# Typification and Plans of Smart-Villages in Rural Area

Yuncheol Nam[1*], Yunyoung Nam[2]
[1]*Department of Architecture, Joongbu University, Goyang, South Korea*
[2]*Department of Computer Science and Engineering, Soonchunhyang University, Asan, South Korea*

*Contact: ycnam@joongbu.ac.kr, phone +82-31-8075-1622

*Abstract*— **There are various issues in rural areas. There are population decline, aging, and the absence of jobs and amenities. The government continues to expand various projects for rural areas. Recently, the government is promoting smart-village projects in farming and fishing villages. The purpose of smart-village is safety, convenience, smartization, and productivity improvement. The purpose of this study is to investigate the project and implementation process of smart-villages, and domestic and foreign cases. And it classifies smart-villages and suggests implementation strategies.**
**The conclusion is as follows.**
**①The smart-village business focuses on safety, living convenience, facility smartization, and agriculture and fisheries. ②In overseas cases, the smart-village project focuses on improving the residential environment of farming and fishing villages in the EU, the UK, and Germany. Japan focuses on improving energy and agricultural and fisheries productivity. ③ It is recommended that the smart-village business be subdivided and promoted as much as possible. And the project enhances synergy in cooperation with other government ministries. ④Smart services increase credibility through FGI for public officials and residents. ⑤The project is carried out in consideration of agricultural products, tourism festivals, natural environment, history and tradition.**

## I. INTRODUCTION

### 1.1 CURRENT STATUS OF SMART-VILLAGE

Currently, various social problems are increasing in rural areas. As people move from rural areas to cities, rural areas are depopulated. In addition, the aging population is increasing, schools are disappearing, and hospitals and cultural facilities are decreasing. The government is implementing various policies targeting rural areas.

As a representative example, a smart-village project in a smart-city is being implemented.

The smart-village business uses ICT to provide safety, living convenience, smart buildings, and smart farms. In particular, the smart-village project aims to improve the rural economy and residential environment by using ICT. This can reduce the population decline in rural areas by narrowing the gap between urban and rural areas.

### 1.2 SCOPE AND CONTENTS OF RESEARCH

The purpose of this study is to review the policy and implementation process of smart-villages promoted by government ministries, domestic and foreign best practices, classify projects, and propose promotion strategies.

The contents of the study are as follows.

(1) Investigate smart-village research terms and major services. (2) Investigate projects and processes of smart-villages. (3) Investigate domestic smart-village business cases. (4) Investigate overseas smart-village business cases. (5) Classify smart-village projects and present action plans.

## Ⅱ. SMART SERVICE ; CONTENT AND TECHNOLOGY

### 2.1 SMART SERVICE CONTENTS

Smart-village has many IT smart services.

Smart-village project is operated with 12 services in 3 fields. The three areas are 'safety improvement', 'life convenience improvement', and 'agricultural productivity improvement'. 'Productivity improvement' mainly means smart farms, and some local governments are also making smart fisheries a task.

The domestic smart-village business has two main characteristics. It can be divided into ① smart-village (safety and convenience) and ② smart farm (smart agriculture).

And rural areas focus on 'agriculture' and carry out smart farm projects. The fishing village carries out the 'smart fishing' business.

### 2.2 REPRESENTATIVE TECHNOLOGY FOR SMART-VILLAGE PROJECT CASES

ICT (Information & Communication Technology) is a key technology for smart-villages. ICTs a compound word of IT (Information Technology) and ICT (Communication technology), and is a concept emphasizing Communication in existing IT. ICT is representative of IoT, AI, AR, and VR technologies, and smart-villages (including smart agriculture and fisheries) mainly utilize IoT and AI.

**Table 1.** Representative Technology for Smart Villages

| SMART TECHNOLOGY | DETAIL | Description |
|---|---|---|
| IoT(Internet of thing) | Sharing information by connecting objects in daily life through wired/wireless networks. | Electronic devices and telecommunications companies (SKT, KT, LGU+) collaborate to provide CCTV, smart home, smart mobility, telemedicine and education, tourism, weather and environmental information, etc.<br>Unmanned drones, unmanned tractors, agricultural equipment, etc. (agriculture and fisheries)<br>Analysis of data such as crop growth status, soil, weather, temperature and humidity, disease and pest prevalence, and provision of optimal guides (agriculture and fishery) |
| AI (Artificial intelligence) | Computer implementation capable of tasks that require human learning, reasoning, and perceptual abilities. | |
| VR(Virtual reality) | Technology that shows (virtual) virtual CG | AR and VR are used for experiencing agricultural activities (how to operate agricultural machinery, grasp crop conditions, and experience harvesting). Compared to IoT and AI, there are still many cases where smart agriculture is mainly applied to games.<br>VR is used for virtual space reproduction, games, etc.<br>AR is to find virtual Pokemon in real life like Pokemon GO. Used in fields such as games, route finding, tourism, and shopping. |
| AR(Augmented Reality) | (Real+Virtual) Technology that synthesizes and overlaps virtual information in real space to show objects or images | |

**Table 2.** Services in Smart Village Project(DOMESTIC

| services | contents |
|---|---|
| residential environment (safety) | Mobility safety control and accident emergency rescue notification (e-Call) service<br>Smart crosswalk<br>'SMART theft prevention system'<br>Radar sensor-based intelligent CCTV<br>Solar Powered Agricultural Black Box<br>IoT sensor of smart farm |
| residential environment (life) | Illegal garbage prevention CCTV<br>Intelligent care service for the elderly<br>Smart parking information and sharing service |
| Smartization of local facilities | 'Smart Senior Citizens' Party + Remote Video System<br><City Hall - Administrative Welfare Center - Senior Citizens' Party> Establishment of welfare service support system<br>'Smart education service of public library' |
| agriculture and fishery | Accident prevention and notification service for agricultural machinery.<br>Unmanned drone control system<br>Autonomous Tractor System<br>Investigate the amount of octopus resources using drones.<br>IoT-based illegal octopus fishing monitoring.<br>Support shellfish aquaculture through drones and CCTV cameras<br>Support pollution prevention system for shellfish aquaculture |

### 2.3 SMART SERVICE CONTENTS IN SMART-VILLAGE PROJECT : SMART-VILLAGE, SMART FARM, SMART FISHERY

In rural areas, there are smart-villages, smart farms, and smart fisheries. ①Smart-villages target residential environments and are divided into safety, living convenience, smartization of local facilities, and productivity improvement (of agriculture and fisheries).

② Smart farms were limited to facility agriculture, but expanded to smart agriculture.

③ Smart fisheries are generally fish farming and drone technology.

## III. SMART- VILLAGE BUSINESS CASE

### 3-1 SERVICES IN SMART VILLAGE (DOMESTIC)

The smart-village project started in 2019, mainly in rural areas. Since 2021, it will be expanded to fishing village smart fishing.

**TABLE 3.** SERVICES IN SMART VILLAGE PROJECT(OVERSEA)

| CASE | DESCRIPTION |
|---|---|
| Eu | Prevent population decline and change<br>Securing budget for public services<br>Suggest linkages between small towns and cities<br>Digitization of rural areas |
| Uk | Resolving rural issues such as segregation of rural areas, population aging, and low productivity<br>Build a digital hub and conduct digital education.<br>In the <E-health Project>, experts such as IT companies + medical teams + universities in the field of medical and welfare work together to build an information system or utilize humanoid robots. |
| Germany | Building a digital ecosystem for rural villages.<br>Operate a platform that provides portal services specialized for rural villages <administrative support, village homepage, local community center, online market, delivery, village news>.<br>'Information Village Project' <E-commerce, establishment of information contents to generate revenue, revitalization of the local economy> is promoted. |
| Japan | Since 2018, we have been focusing on technological innovation in smart agriculture.<br>The goal is to improve agricultural productivity, supplement labor force, and develop advanced agricultural technology by utilizing robot technology and AI.<br>Specifically, autonomous driving of tractors, unmanned weeding robots, and fruit harvesting robots.<br>From 2019, a total system of production-distribution-sales was implemented based on the <balance between food production and demand>. |

3-2 SERVICES IN SMART VILLAGE (OVERSEAS CASE)

As examples of overseas smart-villages, the EU, UK, Germany, and Japan were examined. The main business characteristics in the EU, UK and Germany were concentrated in the residential environment. However, Japan is focusing on agricultural technology. In Fisheries of Japan, 'smart agriculture' focuses on realizing electricity saving, labor saving, and high-quality production by utilizing robot technology or ICT.The characteristic of the smart-village business is that domestic cases provide many services, but foreign cases focus on one or two services. In particular, the case of Japan is implementing smart agriculture in detail.

3-3 SERVICES IN SMART VILLAGE (JAPAN)

The following figure shows the Japanese government's smart-village.

(1) Smart agriculture in Japan

Smart agriculture is defined as realizing new agriculture (unmanned agricultural machinery, agricultural drones, etc.) that realizes energy saving and high-efficiency production by utilizing advanced technologies such as robot technology and ICT (AI, IoT, etc.).

It was also defined as a pattern in which rural areas produce crops and cities consume crops.

①Low-carbon horticultural farming: Development of a system using renewable energy in a vinyl house, a DC drive device, and an energy-saving night temperature controller using a heat pump

② Comprehensive marine aquaculture management: Form a power system by introducing photovoltaic power generation facilities into the farm. Controls the charging and discharging of generated power and develops a wireless remote management system. For example, in an oyster farm, microbubbles are generated to prevent red tide damage and increase productivity.

③ energy management by EMS (Energy Management System): EMS development that is suitable for areas with low energy demand compared to urban areas and realizes low cost through shared use of the area. It has an energy saving effect of about 10%.

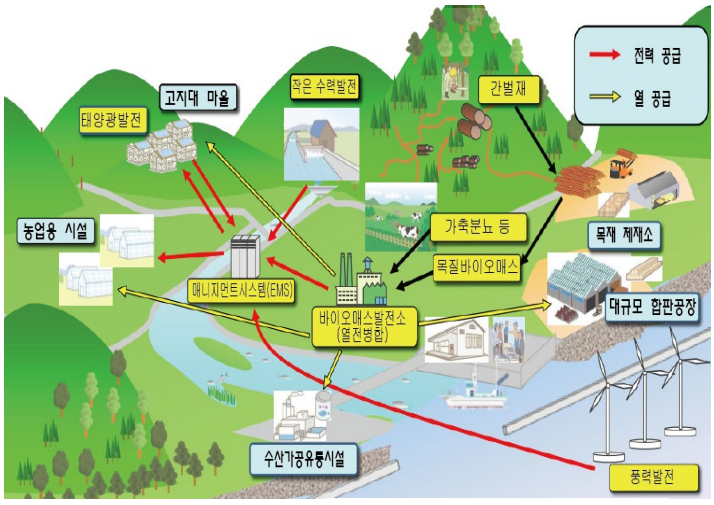



FIG. 1. SMART GRID SYSTEM IN RURAL JAPAN, 2011

https://www.maff.go.jp/j/council/seisaku/nousin/bukai/h23_2/pdf/data3s.pdf, p.17

**Table 4.** Services in Smart Village Project(JAPAN)

| Smart service items | Description |
|---|---|
| 1. Unmanned automatic driving tractor (Iwamesawa City, Hokkaido) | Field plowing is unmanned, and compost and seeding (unmanned + manned) are cooperative. |
| 2. Unmanned automatic driving system for agricultural machinery | Developing safe work technology while remotely moving unmanned agricultural machines to rice fields |
| 3. Automatic operation seeder Kubota Co., Ltd. | Worked from 2 drivers and 2 assistants to unmanned driving + 1 assistant to monitor. |
| 4. Automatic driving seeder and combine (Kubota Co., Ltd.) | Combine with automatic operation assist function |
| 5. Development of a system that remotely and automatically controls water management | The water level of paddy fields is measured with a sensor and transmitted to the cloud, and the water supply valve and water line are remotely and automatically controlled from mobile devices. |
| 6. (Agricultural drone) Development of low-flying remote automatic fertilizer spraying technology (Nile Works Co., Ltd.) | Take-off, dispersal, and landing of drones are automatically performed by button operation on a dedicated tablet. |
| 7. ‘Apare’, a cloud-type farming support service using satellite remote sensors | Establish a web page that analyzes farmland photos taken by artificial satellites to diagnose growth conditions and inform prescriptions. |
| 8. Smart Fertilizer Spreading System (Working Machine for Riding Management Machine) (Iseki Agricultural Machinery Co., Ltd.) | The growth sensor installed in the front measures the condition of the rice and adjusts the fertilizer application amount. The sensor on the rear checks the amount of fertilizer applied and adjusts the amount of fertilizer applied again. |
| 9. Automatic fertilizer system (facility cultivation) tailored to crop growth (Rootrek Network Co., Ltd.) | Control the optimal soil environment by storing insolation, soil moisture, soil temperature, etc. in the cloud and automatically adjusting the supply of liquid fertilizer |
| 10. Agricultural assist suit (Inopice/Power Assistant National) (Inopice, Wakayama University, etc.) | The power of air reduces the burden on the back. Inexpensive, easy to install, no battery required. Useful for heavy work, transportation, etc. |
| 11. Other agricultural technology Remote control weeder (Sanyo Equipment Co., Ltd.), autonomous driving unmanned weeder, support for grape cultivation skills using smart glasses, support for apple pruning technology using 3D and VR, pest diagnosis app using artificial intelligence (AI) Soil disease diagnosis app (diagnosed by image), automatic cabbage harvester, etc. | |

( * Japan MAFF(https://www.maff.go.jp/j/kanbo/smart/))

## IV. Typification and Plans of Smart-Villages

### 4-1 Types of smart village business

In this study, it is classified into the following three types.

① Smart Village (Residential Environment) ② Smart Agriculture ③ Smart Fishery.

In this way, a ‘simple business policy’ can increase the success rate and effect.

And existing rural businesses are based on smart technology. Also, agriculture improves productivity based on smart technology. The fishery improves the productivity of the fishery based on smart technology.

### 4-2 Smart Village Business Plan by Type

(1) 3 type <major classification>: Classified into 3 types ① Smart Village ② Smart Agriculture ③ Smart Fishery.

(2) 5 type <small category>: 3 Smart Village ①-A. Enhance safety, ① -B. Expansion of convenience facilities ① -C. Smartization of local facilities.

② Smart Agriculture and ③ Smart Fisheries aim to improve productivity.

(3) Target area: The scope of the smart village is the residential area of farming and fishing villages.

Smart agriculture is cultivated land (paddy fields, etc.), and smart fishing is the sea.

(4) Main service: The number of smart service items is large. In 2021, it was said that there were about 442 of them.

It is necessary to select appropriate, urgent and important services for rural areas.

(5) Cooperation project: It is important to promote the smart village project in cooperation with other government projects.

(6) Survey and Survey Target: Many residents do not know about the smart village project. Therefore, the interview method, especially the focus group interview (FGI), is a good method.

Also, the role of the relevant public officials in the jurisdiction is important.It is recommended to conduct surveys and interviews with experienced people.

## V. Conclusion

This study investigated cases, types, and plans of smart village projects.

The conclusion is as follows.

(1) Cases of domestic smart villages have been increasing since 2019. Smart services are focused on safety, convenience, facility smartization of residential areas, and smartization of agriculture and fisheries.

(2) Smart villages in the EU, UK, and Germany focus on improving the environment of residential areas in rural areas.

Japan is focusing on rural energy self-sufficiency and smart agricultural machinery.

(3) The types of smart village projects are classified as follows.

① Smart Village (Safety, Convenience, Smartization) ② Smart Agriculture ③ Smart Fishery, etc.

(4) Surveys increase reliability through interviews (FGI).

(5) The smart village project is carried out in consideration of local agricultural products, tourism festivals, and the natural environment.

## REFERENCES


[1] 남윤철, 박은영, 스마트시티사업을 위한 설문결과 분석과 추진방향, 산학기술학회, 2021, vol. 22, no. 3, pp. 422-428

[2] 중부대학교, 포스트미디어, 금산군 스마트도시 조성연구, 2020, p. 107

[3] 한국정보화진흥원, SMART CITY by SMART CITIZEN 제 1 편_시민 주도 스마트시티의 도전과제, 2018 (스마트시티 1 편 https://www.4th-ir.go.kr ? article ? download)

[4] 과기정통부 2022 년 스마트빌리지 지원확대, 전자신문 2021 년 12 월 20 일자 (https://www.etnews.com/20211220000190)

[5] 내년 스마트빌리지 보급·확산사업 '윤곽', 정보통신신문 2022 년 6 월 13 일자 (https://www.koit.co.kr/news/articleView.html?idxno=98417)

[6] 농림수산성(https://www.maff.go.jp/j/kanbo/smart/)

[7] '스마트 농업 육성 및 지원에 관한 법률' 제정(안) (https://www.moleg.go.kr/lawinfo/makingInfo.mo?lawSeq=63619&lawCd=0&&lawType=TYPE5&mid=a10104010000)

[8] 스마트빌리지 서비스 발굴 및 실증사업공모 관련 FAQ (https://www.kica.or.kr/file/download/65e796da-7d81-4780-9db2-eccb8853db5c)

[9] 일본내각부 홈페이지(https://www8.cao.go.jp/cstp/society5_0/smartcity/)

[10] 일본, 비즈니스네트워크 홈페이지 (https://businessnetwork.jp/article/6280)

[11] 일본경제신문 2012 년 6 월 20 일자, 다음은 스마트빌리지가 온다(次は「スマートビレッジ」がやってくる) (https://www.nikkei.com/article/DGXNASFK1804B_Y2A610C1000000/)

# Improved Real-Time Object Tracking on Smartphone

Neunggyu Han[1*] and Yunyoung Nam[2]
[1]*Department of ICT Convergence, Soonchunhyang University, Asan 31538, Korea*
[2]*Department of Computer Science and Engineering, Soonchunhyang University, Asan 31538, Korea*
*Contact: az0422@naver.com

*Abstract*— **Recently, many products and services are being released according to the increase in the use of artificial intelligence. The artificial intelligence smartphone cradle is one of them. The cradle detects and tracks objects and automatically moves to center them. However, the application that control this smartphone cradle have several problems. For example, if two or more objects are overlapping, tracking is stopped. Moreover, objects such as puppies and kittens have inaccurate recognition shapes. And in the previous papers, we implemented an improved application. The application uses YOLO to recognize and track person, dog, and cat. However, the models used in this application also had a problem in that recognition stopped for some objects. In addition, these models also had a problem of slow inference rate. In this paper, we construct new models that solve these problems. These models are based on YOLOv8. In addition, these models show more than 90% accuracy and inference rate of more than 20 fps, showing advanced inference performance compared to previous stages.**

## I. INTRODUCTION

Recently, many products and services are being released according to the increase in the use of artificial intelligence. The artificial intelligence smartphone cradle is one of them. This cradle is mainly used by YouTube creators. This cradle detects and tracks objects and automatically moves them to the center of the smartphone.

The applications that control this smartphone cradle have several problems. For example, if two or more objects are overlapping, tracking is stopped. In addition, when a new object appears, the cradle has a problem of tracking only the new object. Finally, small objects such as puppies and kittens are not detected. In this paper, we implement an application to solve these problems. The application detects and tracks classes of person, dog, and cat.

In previous papers [1], [2], we have implemented an application that detects and tracks an object. In those papers, object detection was successfully, but this application has some problem. For example, small objects such as puppies and kittens were not detected. In addition, this application has more problem that detection stopped when moving away from an object. In this paper, we propose two YOLO-based models such a solved problems by new algorithm.

In this paper, we implement an application using YOLO [3]. YOLO is a representative algorithm in object detection. In addition to this, there are R-CNN [4], Fast R-CNN [5], Faster R-CNN [6], and SSD [7]. However, the R-CNN series is a two-stage model. But YOLO is a single-stage model and has faster inference rate than R-CNN. Therefore, we adopt YOLO for real-time object tracking.

We adopted a model based on YOLOv8 [8] for fast inference and smooth object recognition in a smartphone environment. Unlike other versions, YOLOv8 has fast inference rate. In addition, anchor free is applied to enable smooth object recognition. The anchor is a concept first introduced in YOLOv2 [9] and is a predefined value for the size and coordinates of the bounding box of an object to be inferred. Based on this value, the size of the object is calculated, and inference is made. In YOLOv8, this anchor has been deleted, making it possible to recognize objects more smoothly than the previous version.

First is the accuracy model. This model is a fast model that prioritizes accuracy. This model uses ResNet [10], which uses a residual block, and CSPNet [11], a structure that divides the input into two parts, passes through the network for one part, and concatenates the other part with the result. This CSPResNet structure is also used in YOLOv5 [12] and v8. However, in these models, the bottleneck in which the number of filters is reduced in the middle is not used. In addition, it is featured by faster inference rate than YOLOv8s because the depth of the model is not deep.

Next is the lightweight model. This model is a very fast model that prioritizes fast inference rate. This model is a model in which the number of layers is reduced in the accuracy model. This model is also very fast and has high accuracy.

For these models, based on mAP50, the accuracy model has an accuracy of 93%, and the lightweight model has an accuracy of 90%. Fig.1 is a graphical representation of the mAP50 standard accuracy and computing cost. In addition, the accuracy model based on the smartphone environment has an inference rate of 10 fps, and the lightweight model has an inference rate of 50 fps. In addition, one of the problems in the previous paper, the problem of stopping recognition when moving a little away from an object, has been solved.

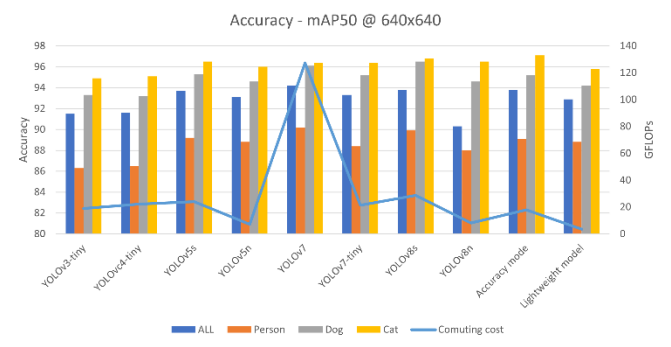



Fig.1 accuracy and computing cost for each model.

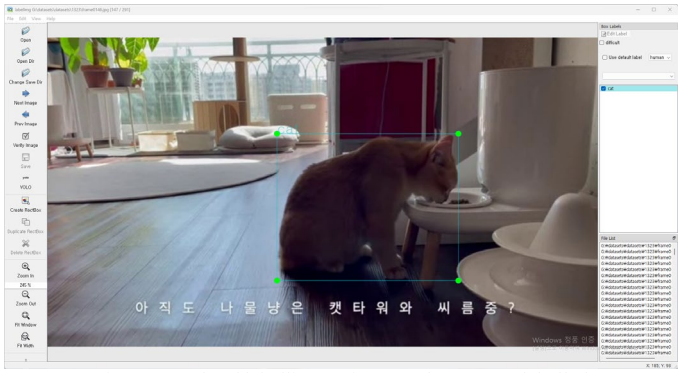



Fig.2 example of labelling on image. The cat was labelled

## II. System Implementation

### A. Environmental Setup

TABLE.1 Training Environment

| | |
|---|---|
| CPU | Intel Xeon Silver 4216 x2 |
| RAM | 192GB |
| GPU | RTX A5000 x2 |
| OS | Ubuntu 20.04 |
| YOLO Platform | Ultralytics YOLOv8 |

TABLE.2 Smartphone Environment

| | |
|---|---|
| Model | GalaxyS21+ |
| OS | Android 13 |
| YOLO Framework | TensorFlow Lite 2.8.0 |

### B. Datasets

We used YouTube video data as our dataset. The reason for this is that it is easy to obtain data with various conditions, for example, composition, angle, brightness, for video data. Therefore, we collected 100 video data. This video data has various resolutions from HD (1280x720) to UHD (3840x2160).

YOLOv8 platform uses images, not videos, for training. Therefore, we extracted frame data in units of 2 fps from video data and used it for training. The extracted frame data was used for training after labeling.

For labeling, automatic labeling was performed on 20 videos using the pre-trained model of YOLOv4-p6 [13]. After going through the verification process, we proceeded with training using the YOLOv7 [14] model to construct a model for labeling. After labeling five videos using this model and going through a verification process, the model was trained again. By repeating this process, labeling was performed on 100 videos and about 65,000 image data.

The labeled data was divided into train and valid sets. 90 video data were used for train, and 10 video data were used for validation.

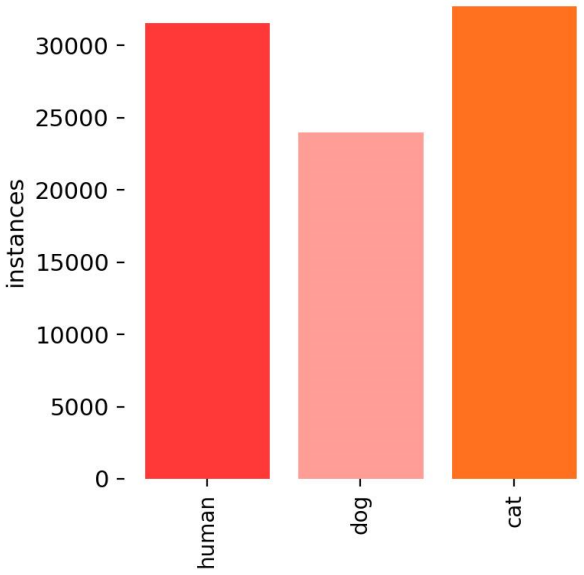



Fig3. the counts of labels for each class for train

### C. Models

Various models were used in this paper. First, YOLOv4-p6 and YOLOv7. This model was used for automatic labeling. Since these models require a lot of computing resources, they were not used in smartphones because they were very slow when used in smartphones.

Next, the two models were used for object recognition in a smartphone. These models were trained on YOLOv8. First is the accuracy model. For this accuracy model, CSPResNet, which applies CSPNet to ResNet, was used. This structure is like the structure used in YOLOv5 and v8, but bottleneck is not used. In addition, a structure cited from CSPNet is used, and consists of a convolutional layer and ResNet structure with down sampling applied on one side, and a structure with max pooling applied on the other side. Fig. 4 shows a diagram of this model.

Final is a lightweight model. The lightweight model is a model formed by reducing the number of layers in the accuracy model. Also, this model has only P3 and P4 as outputs. The reason is to save additional computing resources in the post-processing process. Fig. 5 shows a diagram of this model.

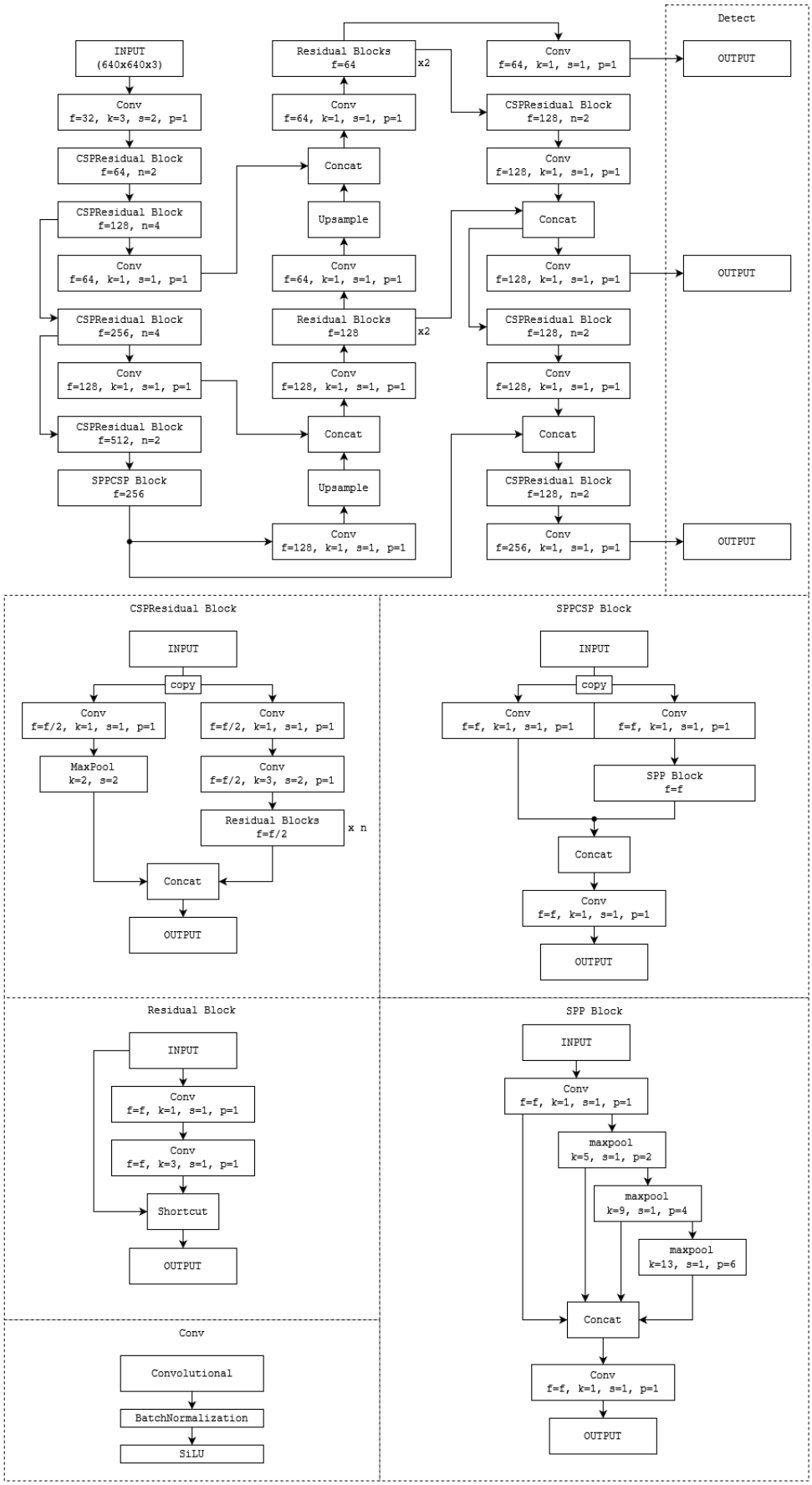



Fig.4 a diagram of accuracy model

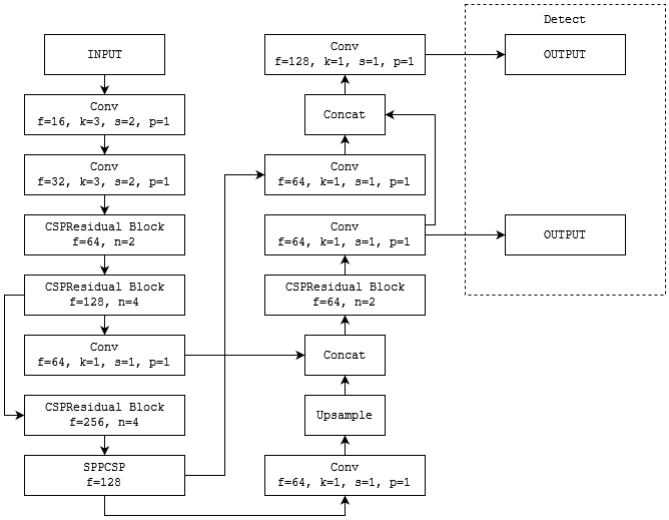



Fig.5 a diagram of lightweight model

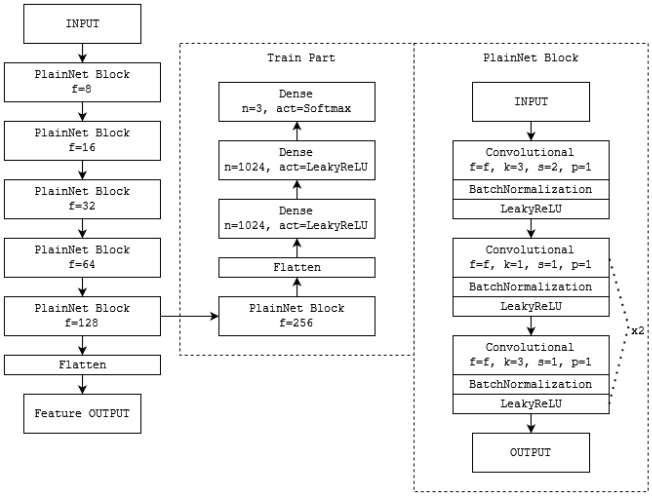



Fig.6 a diagram of CNN of ID assigning system

### *D. ID Assigning System*

TensorFlow Lite only returns coordinates and classes for recognized objects. However, TensorFlow Lite does not provide a system for assigning IDs to object features. Therefore, we need to implement a system that assigns IDs to things.

The system for allocating IDs consists of two steps. The first step is to extract features using CNN. Construct CNN trained on person, dog, and cat, and use them to extract features. A CNN consists of a simple PlainNet. Fig.6 shows the diagram of this CNN.

The system extracts feature and compares the similarity of feature maps to reassign existing IDs or assign new IDs. The formular used here is:

$$Simularity(F_{prev}, F_{curr}) = ReLU\left(\frac{\sum_{i=1}^{n} F_{prev_i} * F_{curr_i}}{\sqrt{\sum_{i=1}^{n} {F_{prev_i}}^2} * \sqrt{\sum_{i=1}^{n} {F_{curr_i}}^2}}\right)$$

Using this formula, if a feature map with a similarity of 50% or more to the existing feature map is found, it is reassigned to the ID of the feature map. Otherwise, a new ID is assigned, and the corresponding feature map is matched with the new ID and stored.

### *E. Application Implementation*

The smartphone application was implemented using TensorFlow Lite. YOLOv8 is basically implemented based on torch. However, torch-based smartphone frameworks have a problem in that inference rate is slow because the use of GPU is not allowed by default. Because TensorFlow Lite allows GPU easily and there are example codes using them, it is easy to implement applications. Therefore, it was implemented using TensorFlow Lite, not torch.

In this application, different input sizes were used for each of the two models. The accuracy model has an input size of 640x640. And the lightweight model has an input size of 320x320.

Fig.7 shows the composition of the object recognition part in this application.

The object recognition screen of this application provides several functions. The first is the ability to select a model. This function can select either an accuracy model or a lightweight model.

Next is the object selection function to be recognized. You can choose to recognize only person, dog and cat, or all objects.

The ID assigned to the object is displayed above the last recognized object. This ID is assigned by analyzing the characteristics of the recognized object, and in the case of the same object, the same ID is assigned.

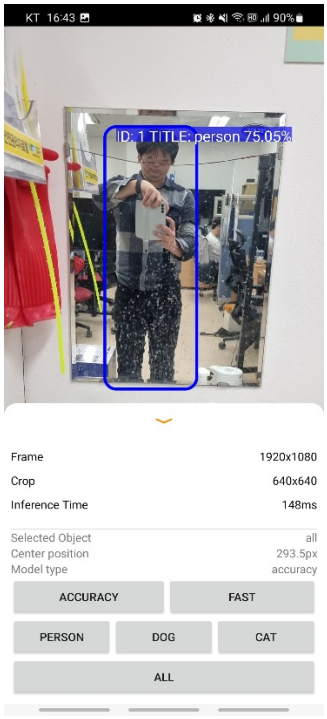



Fig.7 object detection screen of application

## III. Experimental Result

### *A. Model Evaluation*

TABLE.3 shows the evaluation results for each model. And Figs.8-12 is a visualization of accuracy. The models used in this evaluation is not a pre-trained model, but the models trained from our dataset. These models were trained on the YOLOv8 platform. Therefore, anchor free is applied to all models.

In this evaluation, the accuracy was conducted with the validation dataset among our datasets. This dataset contains data from 10 YouTube videos and has about 4,500 images, each labeled with about 450 person, 1,500 dog, and 2,700 cat for each class.

Inference rate also means the inference rate of the model itself. To calculate the accurate inference rate, we proceeded after displaying a black image by covering the camera on the actual smartphone. In addition, after measuring the inference rate for a fair experiment, the smartphone was sufficiently cooled before proceeding. The reason is that the performance of smartphones is degraded due to heat. This is to prevent inaccurate results from occurring.

TABLE.3 evaluation results for each model.

| Model | Input Size | FLOPs (G) | Accuracy (%) | | | | | | | | Inference Rate (fps) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | mAP50 | | | | mAP50-95 | | | | |
| | | | All | Person | Dog | Cat | All | Person | Dog | Cat | |
| YOLOv3-tiny | 640 x 640 | 18.9 | 91.5 | 86.3 | 93.3 | 94.9 | 77.5 | 70.6 | 80.9 | 81.1 | 11.7 |
| YOLOv4-tiny | | 22.1 | 91.6 | 86.5 | 93.2 | 95.1 | 80.5 | 74.7 | 82.4 | 84.5 | 10.9 |
| YOLOv5s | | 24.0 | 93.7 | 89.2 | 95.3 | 96.5 | 85.7 | 79.0 | 89.1 | 88.9 | 10.0 |
| YOLOv5n | | 7.2 | 93.1 | 88.8 | 94.6 | 96.0 | 84.4 | 77.7 | 88.0 | 87.5 | 14.7 |
| YOLOv7 | | 127.3 | 94.2 | 90.2 | 96.1 | 96.4 | 87.2 | 82.4 | 89.1 | 90.1 | 2.78 |
| YOLOv7-tiny | | 21.3 | 93.3 | 88.4 | 95.2 | 96.4 | 85.6 | 79.0 | 89.0 | 88.9 | 10.3 |
| YOLOv8s | | 28.7 | 93.8 | 89.9 | 95.6 | 96.8 | 87.0 | 80.6 | 90.3 | 90.0 | 8.13 |
| YOLOv8n | | 8.2 | 93.0 | 88.0 | 94.6 | 96.5 | 84.4 | 77.7 | 87.8 | 87.8 | 14.9 |
| **Accuracy Model** | | **17.8** | **93.8** | **89.1** | **95.2** | **97.1** | **86.3** | **79.5** | **89.2** | **90.1** | **9.80** |
| **Lightweight Model** | | **3.4** | **92.9** | **88.8** | **94.2** | **95.8** | **83.4** | **76.6** | **87.8** | **85.8** | **20.8** |
| YOLOv3-tiny | 320 x 320 | - | 90.5 | 84.8 | 92.6 | 94.0 | 73.7 | 67.4 | 78.8 | 74.8 | 22.2 |
| YOLOv4-tiny | | | 92.4 | 87.9 | 94.2 | 95.2 | 80.4 | 73.3 | 85.4 | 82.7 | 22.2 |
| YOLOv5s | | | 92.4 | 87.3 | 94.2 | 95.8 | 81.3 | 74.7 | 85.3 | 83.9 | 29.4 |
| YOLOv5n | | | 91.2 | 85.1 | 93.6 | 94.7 | 77.3 | 66.8 | 82.9 | 80.1 | 40.0 |
| YOLOv7 | | | 93.4 | 88.4 | 95.3 | 96.4 | 85.9 | 78.8 | 89.6 | 88.4 | 8.33 |
| YOLOv7-tiny | | | 91.7 | 85.9 | 93.9 | 95.2 | 81.0 | 73.3 | 85.6 | 84.2 | 27.0 |
| YOLOv8s | | | 92.8 | 87.3 | 94.9 | 96.1 | 82.6 | 74.4 | 87.7 | 85.8 | 22.7 |
| YOLOv8n | | | 90.9 | 84.7 | 93.2 | 94.7 | 78.9 | 70.4 | 84.5 | 82.0 | 40.0 |
| **Accuracy Model** | | | **92.5** | **87.5** | **93.9** | **96.1** | **82.6** | **75.2** | **86.2** | **86.2** | **22.7** |
| **Lightweight Model** | | | **90.7** | **85.3** | **92.9** | **94.0** | **74.9** | **66.0** | **82.6** | **76.2** | **50.0** |

*B. Object Detection*

Figs.7-14 show the experimental results in the actual experimental environment. The actual experiment environment is an environment in which real objects are displayed on camera images and then experiments are conducted. This environment proceeded without taking measures to mitigate the performance degradation caused by the smartphone's heat. In addition, additional computing resource consumption due to the ID allocation system is added, resulting in a slightly slower processing rate than the model itself. TABLE.4 summarizes the results of this experiment.

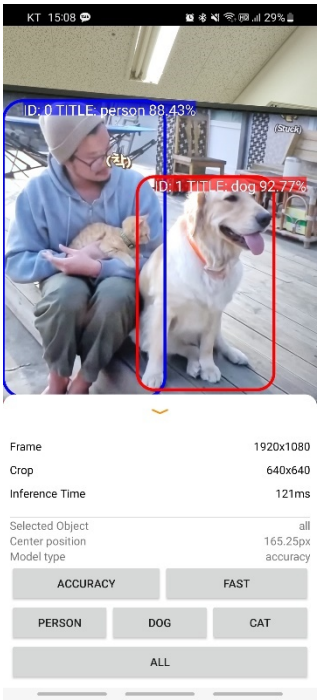



Fig.8 3 classes on smartphone with accuracy model

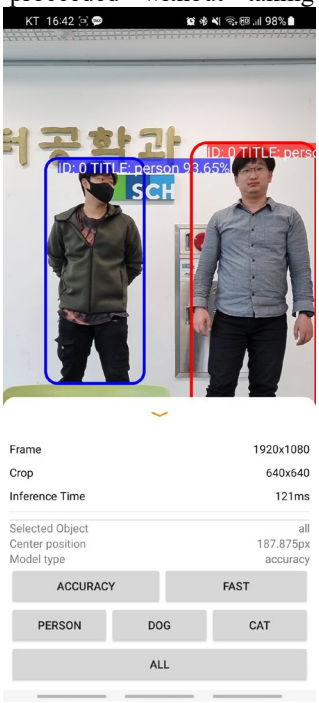



Fig.9 people on smartphone with accuracy model

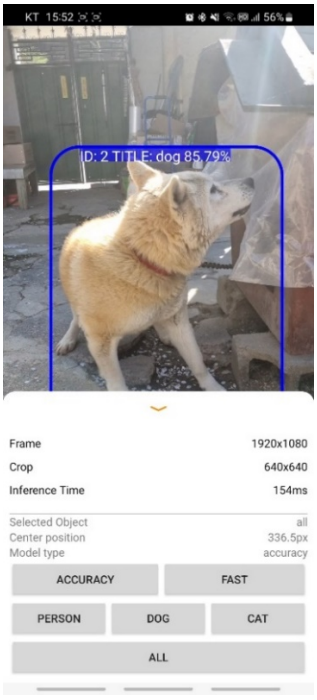



Fig.10 dog on smartphone 9ith accuracy model

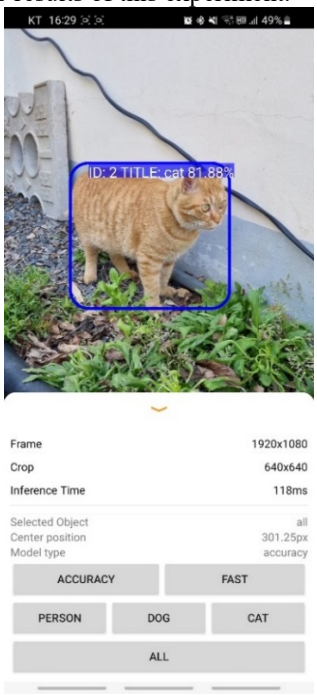



Fig.11 cat on smartphone with accuracy model

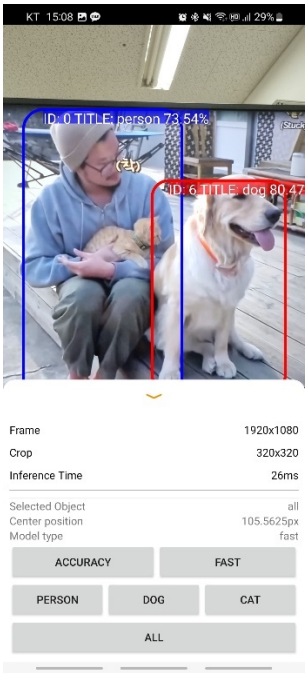



Fig.12 3 classes on smartphone with lightweight model

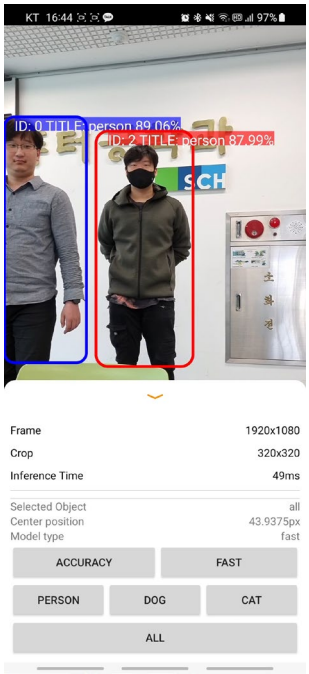



Fig.13 people on smartphone with lightweight model

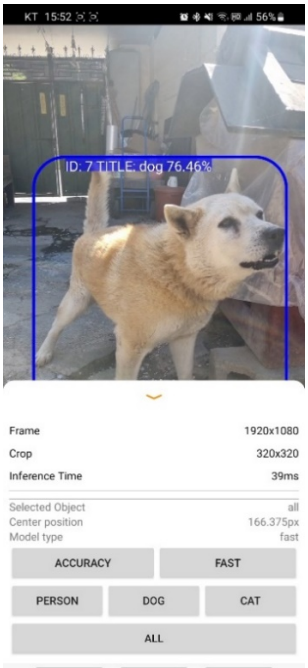



Fig.14 dog on smartphone with lightweight model

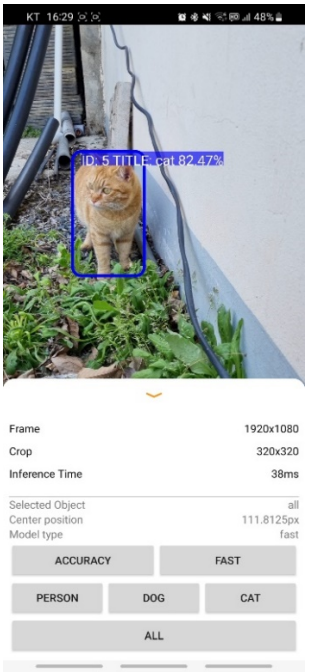



Fig.15 cat on smartphone with lightweight

TABLE.4 summary of object detection results

| Fig. No. | Count of objects | | | Count of detected objects | | | Inference rate (fps) |
|---|---|---|---|---|---|---|---|
| | Person | Dog | Cat | Person | Dog | Cat | |
| Fig.8 | 1 | 1 | 1 | 1 | 1 | 0 | 8.26 |
| Fig.9 | 2 | 0 | 0 | 2 | 0 | 0 | 8.26 |
| Fig.10 | 0 | 1 | 0 | 0 | 1 | 0 | 6.49 |
| Fig.11 | 0 | 0 | 1 | 0 | 0 | 1 | 8.47 |
| Fig.12 | 1 | 1 | 1 | 1 | 1 | 0 | 36.5 |
| Fig.13 | 2 | 0 | 0 | 2 | 0 | 0 | 20.4 |
| Fig.14 | 0 | 1 | 0 | 0 | 1 | 0 | 25.6 |
| Fig.15 | 0 | 0 | 1 | 0 | 0 | 1 | 26.3 |

## IV. Conclusions

In this paper, based on YOLOv8, we proposed two YOLO models that applied CSPNet structure to ResNet. These models are an accuracy model and a lightweight model. In the case of the accuracy model, it showed accuracy like YOLOv8 and had the advantage of requiring fewer computing resources. The lightweight model requires fewer computing resources than YOLOv5n but shows that the accuracy is not significantly lower than other models. In addition, it was shown that real-time performance can be realized to the maximum with inference rates of 10 fps and 50 fps, respectively, when performing inference on a smartphone. If this model is used, it is expected that real-time performance is guaranteed and applications that show smooth recognition performance can be implemented.

However, in the case of the ID allocation system, there is no problem in allocating an ID to a single object. However, when IDs are assigned to multiple objects, there is a problem in that IDs are not accurately assigned according to characteristics. Therefore, there is a need to increase the number of classes for the training data of the CNN model used for feature extraction.

Finally, there is a problem that the inference rate is slow due to the feature extraction model. This is because the structure of the feature extraction model is complex. Therefore, we need to simplify the structure of the feature extraction model.

## Acknowledgment

This research was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0012724, The Competency Development Program for Industry Specialist) and the Soonchunhyang University Research Fund.

## References

[1] Neunggyu Han and Yunyoung Nam, Real-time Object Tracking at Home using YOLO, 2022 Engineering in Biomedical and Rehabilitation Conference

[2] Neunggyu Han and Yunyoung Nam, Real-time Object Tracking Smartphone Cradle Using YOLO, 2022 International Conference on ICT for Smart Health & Home

[3] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[4] Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.

[5] Girshick, Ross. "Fast r-cnn." Proceedings of the IEEE international conference on computer vision. 2015.

[6] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems 28 (2015).

[7] Liu, Wei, et al. "Ssd: Single shot multibox detector." Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016.

[8] Jocher, G., Chaurasia, A., & Qiu, J. (2023). YOLO by Ultralytics (Version 8.0.0) [Computer software]. https://github.com/ultralytics/ultralytics (accessed on 2023.03.20.)

[9] Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[10] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[11] Wang, Chien-Yao, et al. "CSPNet: A new backbone that can enhance learning capability of CNN." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020.

[12] Jocher, G. (2020). YOLOv5 by Ultralytics (Version 7.0) [Computer software]. https://doi.org/10.5281/zenodo.3908559

[13] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. "Scaled-yolov4: Scaling cross stage partial network." Proceedings of the IEEE/cvf conference on computer vision and pattern recognition. 2021.

[14] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." arXiv preprint arXiv:2207.02696 (2022).

# Deep Learning based Sleep Posture Classification

Seungwon Choi[1*], Neunggyu Han[2] and Yunyoung Nam[3]
[1]*Department of Computer Science and Engineering, Soonchunhyang University, Asan 31538, Korea*
[2]*Department of ICT Convergence, Soonchunhyang University, Asan 31538, Korea*
[3]*Department of Computer Science and Engineering, Soonchunhyang University, Asan 31538, Korea*
*Contact: csw011201@sch.ac.kr, phone +82 10-5644-2978

*Abstract*— **Currently, sleep monitoring is performed using a camera or a wearable device such as a smartwatch. In the case of using a wearable device, it may cause inconvenience to the user because the device must be worn on the wrist or the like. Also, when using the camera, it cannot be used if covered with a blanket. Therefore, we propose a new sleep monitoring method. This method uses a pressure sensor and deep learning. In this paper, we use a small pressure sensor to classify 6 classes consisting of 5 postures and blank data. At this stage, more than 90% accuracy was obtained. Therefore, we will use this method to implement a real-time sleep monitoring system that can determine even the respiratory rate using a larger pressure sensor.**

## I. Introduction

We spend as much as a third of our lives sleeping. However, sleep is very important for us to recover from fatigue. That is why sleep is an indispensable activity for us. Therefore, there is a lot of interest in sleep.

Short, Michelle A., et. al. [1] investigated the relationship between sleep and mood in adolescents. Thase, Michael E. [2] studied the relationship between sleep and depression. Patel, Aakash K., et. al. [3] studied sleep stages.

Also, we are trying a lot to get a sound sleep. For example, we monitor our sleep through wearable devices such as smartwatches [4], [5] or use cameras to identify our sleep patterns. [6]-[8] However, these methods may give the user discomfort such as a feeling of restraint because the device must be worn on the body. In addition, in the case of the camera method, there is a risk of leakage of the recorded video due to video recording, and it is impossible to measure if the sleeping person covers the blanket.

Therefore, we would like to propose a new sleep monitoring method. This method does not use a camera or a wearable device such as a smartwatch but monitors by installing a large sensor on the bed. This method is a method of monitoring the user's sleep by installing a pressure sensor on a mat used in a bed to determine the user's lying posture, direction, location, respiratory rate, and the like. Since the user does not wear a separate device, the user does not feel constrained and there is no risk of privacy exposure or interference from blankets.

In this paper, we implemented a sleep posture classifier based on the user lay down using fabric pressure sensor and deep learning technology. We made it possible to distinguish a total of 6 classes in the case of 5 postures and 1 blank data. Each posture includes recumbent to the right, recumbent to the left, prone to the right, prone to the left, and supine position. We collected a total of 10500 data from 6 people. And we constructed a model based on ResNet[9] and trained it with the collected data.

Traditional CNN models include LeNet [10], AlexNet [11], GoogLeNet [12], and ResNet. LeNet is suitable for simple datasets such as handwritten digit recognition. AlexNet is suitable for large-scale image classification problems. GoogLeNet can reduce the number of parameters efficiently with high performance. However, the disadvantage is that the structure is relatively complex. ResNet shows good performance even in deep network structures and is suitable for large-scale image classification problems. It is also relatively simple to write. In this paper, we conducted training with a simple dataset. However, future studies will use more complex and larger dataset. Therefore, we created a model based on ResNet, considering that the data changes on a large scale and the depth of the model deepens.

This model was able to obtain results with an accuracy of more than 94%. Therefore, we plan to utilize this method and bigger pressure sensor in future research.

## II. System Implementation

### A. Environmental setup

TABLE.1 Environmental setup

| | |
|---|---|
| **OS** | Ubuntu 20.04 |
| **GPU** | RTX 3080 |
| **CPU** | i7-12700K |
| **RAM** | 32 GB |
| **Sensor** | Fabric Pressure Sensor(70x70(cm)) |

### B. Pre-processing

Firstly, we collected data and proceeded with pre-processing. A fabric pressure sensor and an MCU were used. In the case of the fabric pressure sensor, the size is about 70x70 (cm). And we were able to get the data of the

sensor with the size of 11x11. Figs. 1-2 shows the fabric pressure sensor and MCU.

Fig. 1 fabric pressure sensor

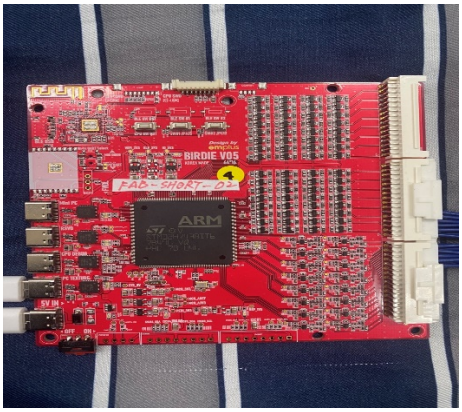



Fig.2 MCU

We connected the devices to a computer to collect data. We collected a total of 10,500 data from 6 people. We collected 300 data from each of the 5 classes per person. And we also collected 1,500 data even when nothing was placed on the pressure sensor. We stored and normalized the collected data. Then, the data entered in 11x11 was processed as an image and resized to 32x32 to facilitate train.

Second, we performed labeling for each data. It was labeled as 0 if it is blank, 1 if recumbent to the right, 2 if prone to the left, 3 if recumbent to the left, 4 if prone to the right, and 5 if lying down in a supine position. Figs. 3-8 show examples of sleeping posture and labeling of data for each class.

Finally, we used the data from 5 people to train and data from one person for validation.

### C. *Model*

We built a model based on ResNet. The reason we used ResNet was that it can stably use for learning, even if the depth of the network deepened in subsequent studies. In general, there is a problem that the error increases as the number of parameters increases as the neural network deepens. ResNet is a neural network that came up with a residual learning method to solve this problem. The input size of this model is 32x32. Fig. 9 shows a diagram of this model.


Fig. 3 blank


Fig.4 right recumbent


Fig. 5 left prone


Fig. 6 left recumbent


Fig. 7 right prone


Fig. 8 supine

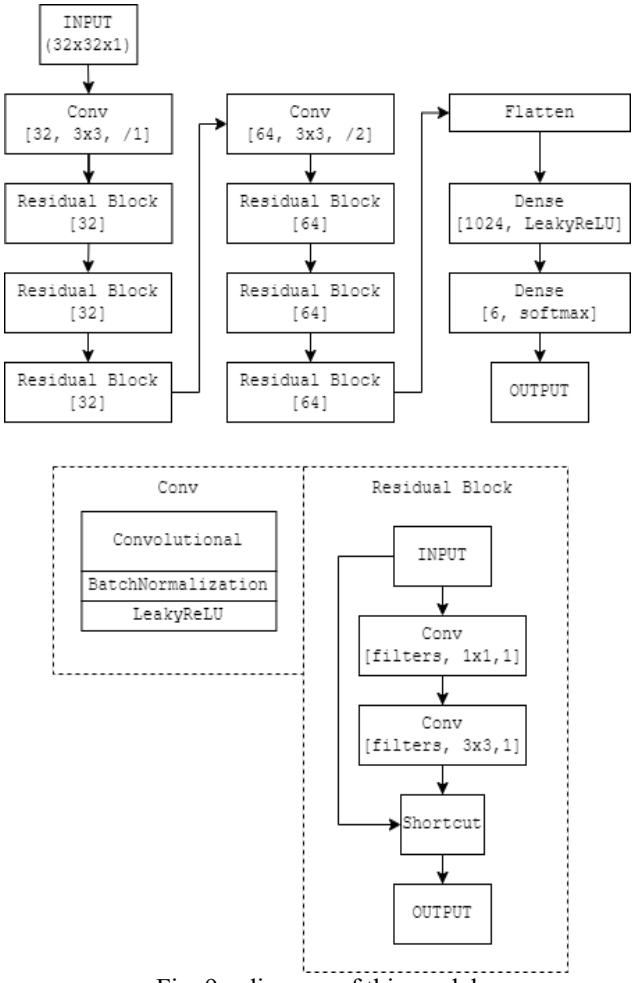



Fig. 9 a diagram of this model

## III. EXPERIMENTAL RESULT

As shown in Fig. 11, even in the case of blank data, some noise is captured. However, this model achieved an accuracy of over 94% even when the noise is present. Fig. 10 shows the confusion matrix for the verified results. And Figs. 11–16 show examples of verified results.

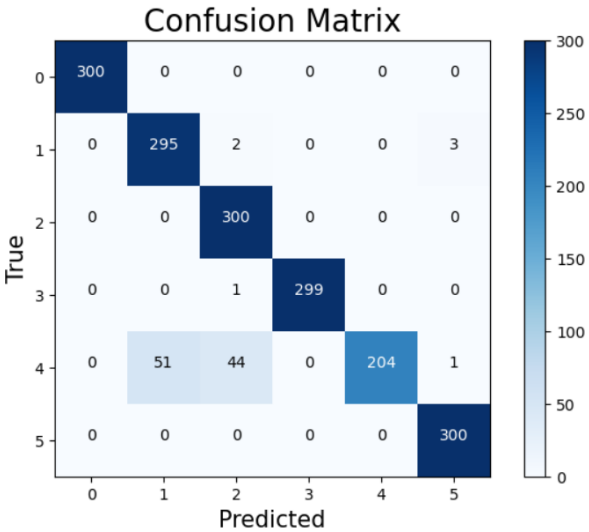



Fig. 10 a Confusion Matrix

```
predictions = model.predict(tresult) #prediction
pred = str(np.argmax(predictions[200])) #prediction result
rlabel = str(tlabel[200]) #labeling of test data
plt.title("Prediction =  "+pred+" Labeling = "+rlabel)
plt.imshow(tresult[200],'gray') #test data
plt.show()
```

```
57/57 [==============================] - 0s 7ms/step
```

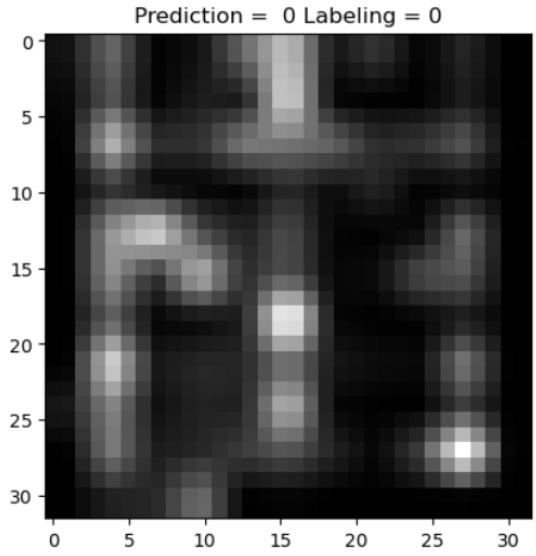



Fig. 11 validation result of blank

```
predictions = model.predict(tresult) #prediction
pred = str(np.argmax(predictions[400])) #prediction result
rlabel = str(tlabel[400]) #labeling of test data
plt.title("Prediction =  "+pred+" Labeling = "+rlabel)
plt.imshow(tresult[400],'gray') #test data
plt.show()
```

```
57/57 [==============================] - 0s 7ms/step
```

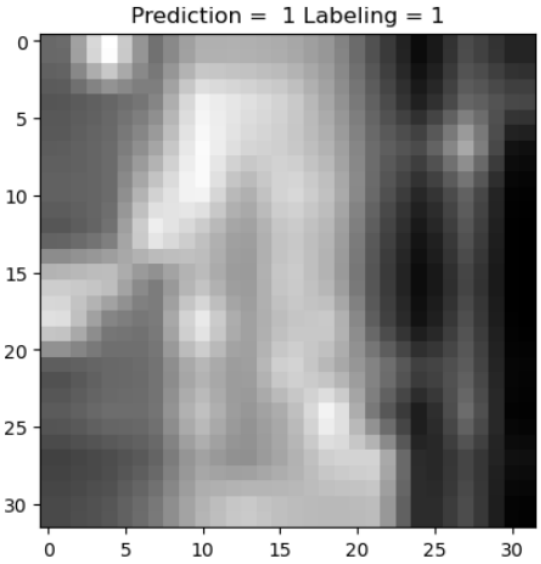



Fig. 12 validation result of right recumbent

```
predictions = model.predict(tresult) #prediction
pred = str(np.argmax(predictions[750])) #prediction resu
rlabel = str(tlabel[750]) #labeling of test data
plt.title("Prediction =  "+pred+" Labeling = "+rlabel)
plt.imshow(tresult[750],'gray') #test data
plt.show()
```

```
57/57 [==============================] - 0s 7ms/step
```

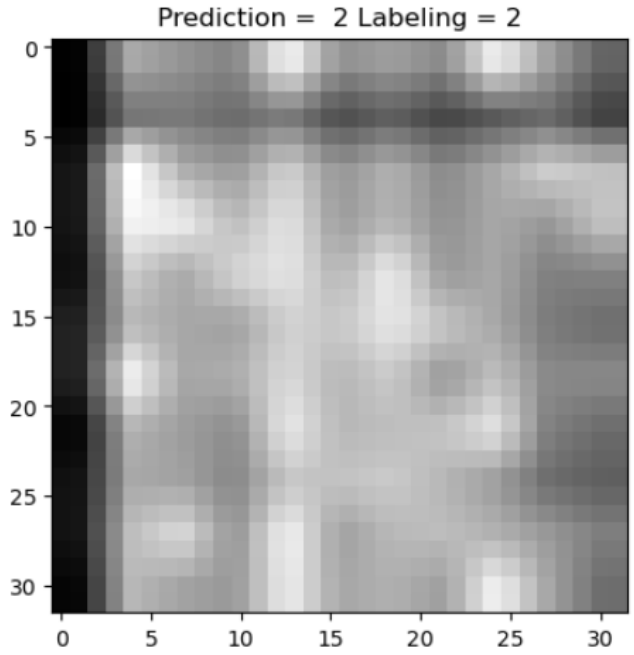



Fig. 13 validation result of left prone

```
predictions = model.predict(tresult) #prediction
pred = str(np.argmax(predictions[900])) #prediction result
rlabel = str(tlabel[900]) #labeling of test data
plt.title("Prediction =  "+pred+" Labeling = "+rlabel)
plt.imshow(tresult[900],'gray') #test data
plt.show()
```

```
57/57 [==============================] - 0s 6ms/step
```

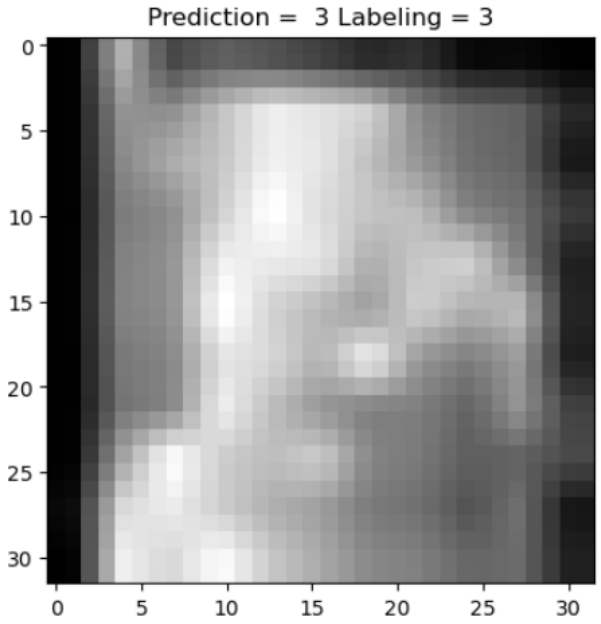



Fig. 14 validation result of left recumbent

```
predictions = model.predict(tresult) #prediction
pred = str(np.argmax(predictions[1200])) #prediction result
rlabel = str(tlabel[1200]) #labeling of test data
plt.title("Prediction =  "+pred+" Labeling = "+rlabel)
plt.imshow(tresult[1200],'gray') #test data
plt.show()
```

```
57/57 [==============================] - 0s 8ms/step
```

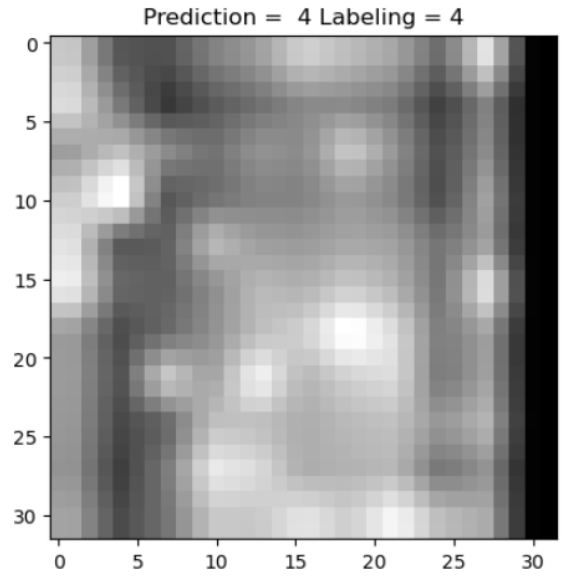



Fig. 15 validation result of right prone

```
predictions = model.predict(tresult) #prediction
pred = str(np.argmax(predictions[1520])) #prediction result
rlabel = str(tlabel[1520]) #labeling of test data
plt.title("Prediction =  "+pred+" Labeling = "+rlabel)
plt.imshow(tresult[1520],'gray') #test data
plt.show()
```

```
57/57 [==============================] - 0s 6ms/step
```

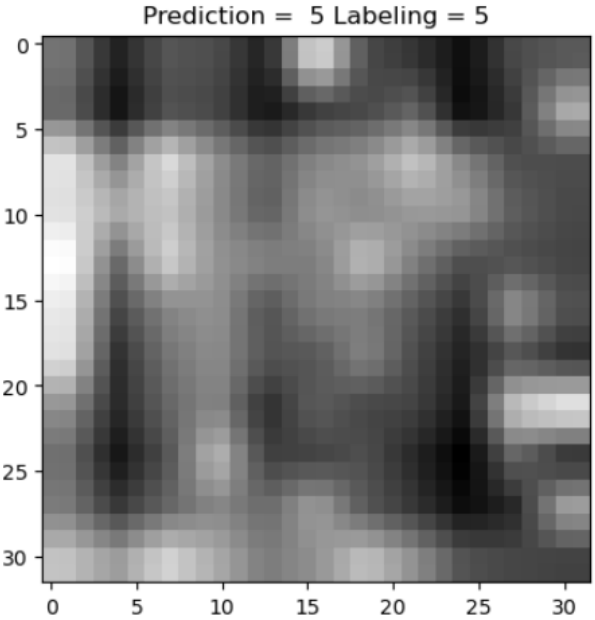



Fig. 16 validation result of supine

## IV. Conclusions

In this paper, we propose a new sleep monitoring method. This method uses a pressure sensor. Unlike wearable devices such as smartwatches, it does not cause any inconvenience to the user during the sleep monitoring process. And, unlike camera footage, this method has no interference from blankets and no risk of privacy exposure. In this study, we developed a ResNet-based classification model that classifies data of the sleep posture coming from the pressure sensor. The result of the model was able to obtain 94% accuracy. We used ResNet instead of other traditional CNN models such as LeNet, AlexNet, and GoogLeNet. Because ResNet can create a model relatively simply, and the size of the model could be increases when the model is deep, but it can be used stably. It is also suitable for large-scale image classification problems. Since our model was created based on ResNet, it can be used with stability even if there is a lot of data and complexity or the depth of the model deepens in future research. Based on this research, we plan to implement a system that classifies actual bed sensor data and sleep monitoring. Currently, we can only classify the sleep posture using the pressure sensor, but in the future, we plan to implement a real-time sleep monitoring system that can detect the user's position on the bed, respiratory rate, sleep time, and more, using a larger pressure sensor.

## Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2022H1D8A3038040)

## References

[1] Short, Michelle A., et al. "The relationship between sleep duration and mood in adolescents: A systematic review and meta-analysis." *Sleep medicine reviews* 52 (2020): 101311.

[2] Thase, Michael E. "Depression and sleep: pathophysiology and treatment." *Dialogues in clinical neuroscience* (2022).

[3] Patel, Aakash K., Vamsi Reddy, and John F. Araujo. "Physiology, sleep stages." *StatPearls [Internet]*. StatPearls Publishing, 2022.

[4] Kwon, Shinjae, Hojoong Kim, and Woon-Hong Yeo. "Recent advances in wearable sensors and portable electronics for sleep monitoring." *Iscience* 24.5 (2021): 102461.

[5] Halson, Shona L. "Sleep monitoring in athletes: motivation, methods, miscalculations and why it matters." *Sports medicine* 49.10 (2019): 1487-1497.

[6] Nochino, Teruaki, et al. "Sleep stage estimation method using a camera for home use." *Biomedical engineering letters* 9 (2019): 257-265.

[7] Masek, Martin, et al. "Sleep monitor: A tool for monitoring and categorical scoring of lying position using 3D camera data." *SoftwareX* 7 (2018): 341-346.

[8] Deng, Fei, et al. "Design and implementation of a noncontact sleep monitoring system using infrared cameras and motion sensor." *IEEE Transactions on Instrumentation and Measurement* 67.7 (2018): 1555-1563.

[9] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[10] LeCun, Yann, et al. "Gradient-based learning applied to document recognition." *Proceedings of the IEEE* 86.11 (1998): 2278-2324.

[11] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Communications of the ACM* 60.6 (2017): 84-90.

[12] Szegedy, Christian, et al. "Going deeper with convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

[13] "Force Sensitive Resistor (FSR)", adafruit, last modified Dec 1, 2022, accessed April 11, 2023, https://cdn-learn.adafruit.com/downloads/pdf/force-sensitive-resistor-fsr.pdf

# Effects of robot-assisted left-hand training on hemispatial neglect in older patients with chronic stoke

Si-An Lee[1], Seong A Lee[2], Jin-Hyuck Park[2*]

[1]*Department of ICT convergence, The Graduate School, Soonchunhyang University, Asan, Republic of Korea*
[2]*Department of Occupational Therapy, College of Medical Science, Soonchunhyang University, Asan, Republic of Korea*
*Contact: jhpark1217@sch.ac.kr, phone +82-41-530-4773

*Abstract*—**Despite of a variety of rehabilitative technique identified to ameliorate neglect symptoms of patients with stoke, the effects of limb activation using a robotic device are still unknown. The purpose of this study was to investigate the effects of the robot-assisted hand training on hemispatial neglect of older patients with chronic stroke. The participants were randomly allocated to the experimental group receiving robot-assisted left-hand training (n = 12) or the control group receiving conventional treatments for neglect symptoms (n = 12). All participants received 20 sessions for 4-week. To investigate training's effects, the Albert's test, and the Catherine Bergego Scale (CBS) were utilized at pre-intervention and post-intervention. After the 20 training sessions, improvements in the Albert's test, and the CBS were found in the EG. However, there was no significant improvement in the Albert's test in the CG. In addition, the EG achieved a greater gain in all outcome measures compared to the CG ($p$<.05). These findings indicate that robot-assisted left hand training was clinically helpful in improving hemispatial neglect in the elderly with chronic stroke compared to the conventional occupational therapy. Robot-assisted left limb activation could be used as an alternative treatment for hemispatial neglect in the elderly with chronic stroke.**

## I. INTRODUCTION

Hemispatial neglect is a common symptom after right hemispheric stroke, and 13 to 82% of patients stroke experience it [1]. Hemispatial neglect, which has a poor effect on the rehabilitation prognosis of stroke patients, is characterized by a diminished or total loss of awareness and reaction to stimuli on the side of the body or space contralateral to a brain injury [2].

There are a number of rehabilitation methods that have been found to be clinically effective in reducing hemispatial neglect, including visual scanning, prism adaptation, transcutaneous electrical neural stimulation technique, and limb activation [3-6].

In some of these researches, a limb activation therapy was employed to lessen visuospatial deficits in both acute and chronic stages of hemispatial neglect [2,7-9]. Furthermore, both passive and active left hand movements are associated a decrease in hemispatial neglect [10]. Given the high incidence of upper limb paresis associated with the neglect syndrome, the positive effect of passive or active limb movement on hemispatial neglect could gain a lot of attention in rehabilitative field [9,11].

Due to the close relationship between motor circuits and attention, which improves the conscious perception of the contralateral hemisphere, the premotor theory of spatial attention proposes that somatosensory activation in the contralateral side by activating the limb facilitates the neural circuits underlying space representation [3,11,12]. Thus, activating the damaged hemisphere via the contralateral limb may lessen the inhibitory competition from the intact hemisphere [3,9,13].

There is an increasing interest in adopting robotic devices in rehabilitation as they could give care more consistently and precisely than therapists [14]. Indeed, during the past ten years, a number of robotic devices have been created to improve paretic upper limbs [15]. Few research has examined the impact of robot-assisted therapy on hemispatial neglect, as opposed to upper limb motor function. [3]. The lack of control condition made it difficult to establish the effects of robot-assisted therapy on hemispatial neglect, even though the prior case series studies indicated significant effects of robot-assisted limb activation on hemispatial neglect. Additionally, since most of subjects are patients with acute stroke in prior studies [16,17], it is challenging to exclude effects of spontaneous recovery on hemispatial neglect.

In order to determine the effects of robot-assisted hand training on hemispatial neglect in older adults with chronic stroke, this study with a randomized controlled trial design was employed. This study proposed that hemispatial neglect would improve more in robot-assisted hand therapy than in conventional occupational therapy.

## II. METHODS

### A. Design

This study was a randomized controlled trial design. The experimental group (EG) or the control group (CG) were randomly assigned to all participants using a random number generated by an occupational therapist who was blind to the allocation using computer software (Excel, Microsoft). The CG participated in traditional neglect therapies, including both remedial and compensatory techniques, such as vibration on the left neck extensors and robot-assisted hand training for the EG. An assessor with more than five years of clinical expertise who was blinded to the group allocation looked at the outcomes at pre- and post-intervention. The intervention consisted of 20 sessions conducted five times a week for four weeks. This study was approved by the local Institutional Review Board. All

participants provided written informed consent before the study according to the code of ethics of the World Medical Association (Declaration of Helsinki, version 2004).

### *B. Participants*

A total of 24 chronic stroke patients with hemispatial neglect were recruited by the author from a rehabilitation hospital in South Korea. The inclusion criteria derived from the previous study (Varalta et al., 2014) were: (1) over 65 years of age, (2) right hemisphere stroke confirmed by a computed tomography scan or magnetic resonance imaging, (3) first-ever ischemic or hemorrhage stroke, (4) intact global cognitive function confirmed by the Korean version of Mini-Mental State Examination score ≥ 24, (5) time since stroke onset ≥ 6 months, and (6) the presence of hemispatial neglect diagnosed by performance on the line bisection test. The exclusion criteria were: (1) left upper limb sensory deficit or impairment, (2) visual impairment, (3) the modified Ashworth scale score for left-hand muscle tone ≥ 2, (4) orthopedic conditions involving the left upper limb, and (5) apraxia [3].

The sample size calculation was conducted by using G*Power 3.1.3 (Informer Technologies, Dusseldorf, Germany). According to a previous study, the effect size was set at 1.31, the α error at a probability of 0.05, and the power at 0.90 [17], resulting in a minimum of nine subjects required in each group. Given a 30% dropout rate, 12 subjects were included in each group.

### *C. Participants*

The EG used the Amadeo® Robotic device (Trymotion GmbH, Graz, Austria) to conduct 20 sessions (five days per week for four weeks) of robot-assisted hand training (Figure 1). A passive rotational joint is positioned between the fingertip and an entity moving laterally on the end-effector-based Amadeo® Robot, which has five degrees of freedom (DOFs). The thumb has two passive rotational joints. The workspace of the fingers is almost totally covered by the five separate translational DOFs. Elastic bands or plasters are used to create the interface between the human hand and the machine, and a Velcro strap is used to restrict wrist movement.

There were four 30-minute sessions. According to a prior study, the exercises were performed as follows: (1) grab and release training (15 minutes of digital joint flexion/extension exercises from the thumb to the fifth finger); and (2) count training (15 minutes of counting a number sequence from one to five). Through the assistive treatment mode of the Amadeo® robot, the participant's hand movements were helped by the robot and modified to the individual's degree of function. The EG participants were given visual feedback of their hand motions during the training via video animation displayed on a monitor.

Conventional occupational therapy for the symptoms of hemispatial neglect was administered to the CG over the course of 20 sessions, each lasting 30 minutes. It included vibration stimulation administered to the middle of the left forearm and the left neck extensors, as well as visual scanning training using a prism. Additionally, the CG participants learnt how to tilt their head or trunk as a compensatory technique for easing hemispatial neglect symptoms. All sessions were run by two occupational therapists who were dependent and had more than five years of experience.

Fig. 1. Amadeo® robotic device

### *D. Outcome measures*

The training effects were investigated by using the Albert's Test, and the Catherine Bergego Scale (CBS). The Albert's Test required the subject to mark 40 lines that were distributed at random on a test sheet. The demonstrator for the assessor marked four lines in the center. As a result, 36 lines were determined to be missing [18, 19].

The Catherine Bergego Scale (CBS) is a checklist with a total score of 30 that is used to evaluate the signs of neglect by observing and interviewing the participant's everyday functioning in 10 actual circumstances. A 4-point scale is used to evaluate each item. A subject who receives a total score of 0 is deemed to not have any neglect syndrome effects on daily life activities. More severe neglect symptoms' effects on daily living are indicated by higher scores [20]. The CBS is a technique that evaluates hemispatial neglect symptoms as well

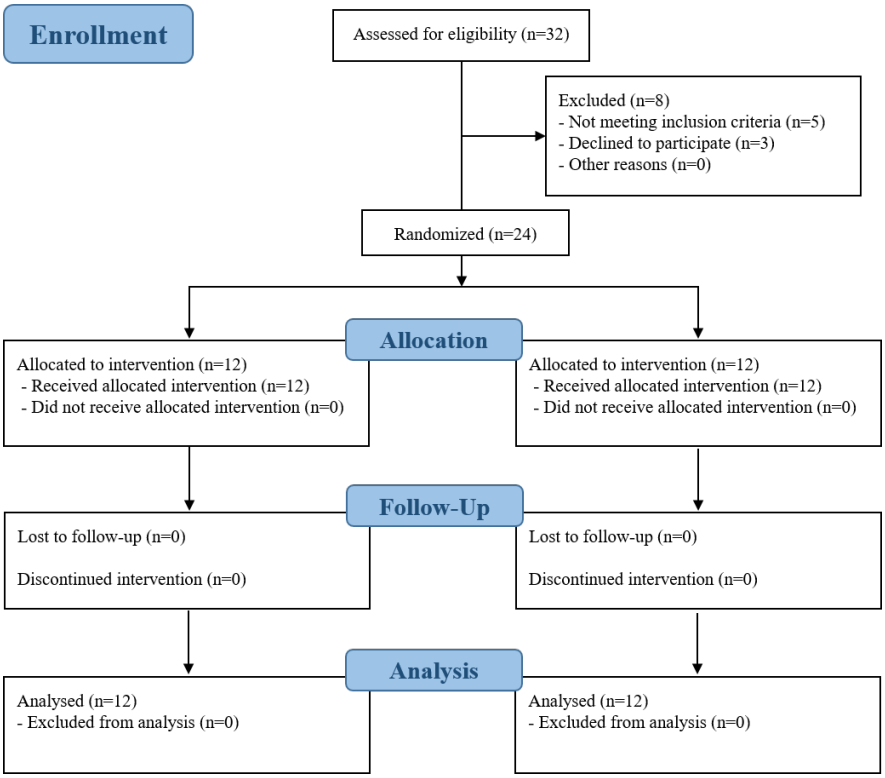



Fig. 2. Flow diagram of the study process

as the degree of disease recognition by comparing scores from observation and interview, unlike the the Albert's test [21].

### *E. Statistical analyses*

All data were analysed using SPSS 22.0 version. All measures were described as the mean ± standard deviation. The Shapiro-Wilk test was used to determine normal distribution of outcome measures. To compare the general characteristics of the participants between both groups, the Chi-square test and independent t-test were conducted. After the training sessions, the differences between the groups were analysed by using the repeated-measure analysis of variance (ANOVA). The effect size of each intervention group was calculated using the partial η2 value. A partial η2 ≥ 0.14 was considered a large effect; between ≥ 0.06 and <0.14, a moderate effect; and a between ≥ 0.01 and < 0.06, a small effect [22]. The statistical significance level was set at $p < 0.05$.

## III. RESULTS

### *A. General characteristics of the participants*

A total of 24 participants were selected from 32 stroke patients. All participants completed the 20 training sessions. No dropouts were recorded during the intervention and all participants fulfilled the intervention protocol (Figure 2). There were no significant differences in the general characteristics between both groups (Table I).

TABLE I

DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF BOTH GROUPS

| Characteristics | | Experimental group (n = 12) | Control group (n = 12) | $\chi^2$ / t |
|---|---|---|---|---|
| Demographic characteristics | | | | |
| Sex | Male | 7 (53.8%) | 6 (50.0%) | .168 |
| | Female | 5 (41.7%) | 6 (50.0%) | |
| Age (years) | | 69.08 ± 4.71 | 71.58 ± 3.17 | 1.523 |
| Stroke type | Infarction | 6 (50.0%) | 7 (58.3%) | .168 |
| | Hemorrhage | 6 (50.0%) | 5 (41.7%) | |
| Onset period (years) | | 9.50 ± 2.61 | 9.08 ± 2.10 | .430 |
| Clinical characteristics | | | | |
| LBT (mm) | | 9.38 ± 1.29 | 8.66 ± 1.13 | 1.447 |

LBT, Line Bisection Test

### *B. Performance on hemispatial neglect tasks*

TABLE II

COMPARISON OF OUTCOME MEASURES IN BOTH GROUPS

| | Experimental group (n = 12) | Control group (n = 12) | Between-group differences | F | $\eta^2$ |
|---|---|---|---|---|---|
| Albert's test (score) | | | | | |
| Pre-intervention | 7.67 ± 2.29 | 6.75 ± 1.60 | | | |
| Post-intervention | 5.00 ± 1.34 | 5.92 ± 1.56 | 1.833 (0.23; 3.43)* | 5.664* | .205 |
| Within-group changes | 2.66 ± 1.96 (1.41; 3.91) ** | 0.83 ± 1.80 (-0.31; 1.97) | | | |
| CBS (score) | | | | | |
| Pre-intervention | 20.75 ± 2.76 | 19.75 ± 1.86 | | | |
| Post-intervention | 15.83 ± 3.73 | 18.50 ± 1.31 | 3.666 (2.25; 5.07)*** | 29.014*** | .569 |
| Within-group changes | 4.30 ± 1.76 (3.03; 5.56) *** | 1.25 ± 1.42 (0.34; 2.15) * | | | |

Repeated-measures ANOVA indicated that there were significant group × time interactions for the Albert's test ($p < 0.05$; $\eta2 = 0.205$), suggesting that robot-assisted hand training showed a greater improvement in ameliorating hemispatial neglect symptoms compared to the conventional treatment (Table II).

### *C. Hemispatial neglect symptoms in activities of daily living*

There was a significant group × time interaction for the CBS ($p < 0.001$; $\eta2 = 0.569$). According to the current study, robot-assisted hand training was more clinically effective in easing symptoms of hemispatial neglect in the participants' everyday activities (Table II).

## IV. DISCUSSION

The main purpose of this study was to investigate the effects of limb activation using a robotic device on hemispatial neglect in older adults with chronic stroke. According to the study's findings, robot-assisted left-hand training may enhance visuospatial exploration as determined by Albert's test, as well as functional performance in daily living as determined by the CBS. This suggests that robot-assisted hand training may be helpful in reducing the neglect symptoms experienced by older adults with chronic stroke.

These results were consistent with previous finding on the effectiveness of contralateral limb activation for reducing the level of impairment in stroke patients with hemispatial neglect due to stroke [3,9-11]. The findings of the current investigation are supported by earlier research that suggested passive and active left finger movements could lessen neglect relative to visual cueing in the neglect sides [7,13]. Similar to this, it was established in the earlier work that proprioceptive cueing regarding the position of the left finger might improve the hampered spatial representation of the left side [11].

For stroke patients with neglect symptoms, limb activation using robotic devices has been attempted in earlier trials. Three stroke patients' hemispatial neglect of the left hand was treated using robot-assisted hand training, according to reports [3,9]. These results agreed with those of the current study, despite the fact that the latter's design was a little different. The patients in the prior research had acute strokes, but this study verified the effects of robot-assisted limb activation on older patients with chronic strokes who had hemispatial neglect, allowing this study to rule out the effects of spontaneous recovery. Additionally, more data was gathered in this study's randomized controlled trial approach than in a case-series design or a single-subject design [3,9].

The benefits of robot-assisted hand training may be explained by the fact that using left-handed motions to activate right hemisphere motor circuits causes the related attentional mechanism to be recruited, which enhances attention for the left side of the area [5]. In fact, Takahashi and colleagues (2008) discovered using functional magnetic resonance imaging that robot-assisted hand training increased the size of the primary sensorimotor cortex representational map [23]. Furthermore, seeing an animation of their hand motions on the monitor could have caused the participants in the EG to become more aware of their extrapersonal space. As a result, the results of this study imply that left-hand training with a robot in the left space may improve the neural networks supporting spatial representation in the injured right hemisphere. However, from a rehabilitation standpoint, robot-assisted limb activation may also be employed to enhance upper limb motor function. In fact, robot-assisted limb activation enhanced gross motor function and neglect symptoms of the left upper limb in the prior study [3]. When considered as a whole, robot-assisted limb activation is believed to be clinically advantageous for stroke patients who exhibit neglect symptoms.

Some elements of this study were more beneficial than those in an earlier trial. The effects of robot-assisted limb activation on a participant's daily function were not examined in the prior

study [3]. The improvements that were confirmed in the CBS, however, suggest that the training effects are only partially generalized to daily life in the study's findings. Additionally, compared to paper and pencil examinations like the Albert's test, behavioral assessments may be more sensitive to the presence of neglect [21]. These findings revealed the clinical significance given that maintaining independence in daily life activities is one of the primary goals of rehabilitation.

The very limited sample size of this study limits generalizing the benefits of robot-assisted left-hand training to all stroke patients, even though it was able to support the applied treatment idea. It is also unclear how much right hemisphere activation is useful for hemispatial neglect because this study did not evaluate right hemisphere activation using neuroimaging devices. Therefore, before this method is widely used in clinical settings, subsequent research need to demonstrate that robot-assisted limb activation works with bigger study samples and the incorporation of neuroimaging devices.

## V. Conclusion

The outcomes of this study demonstrated that the hemispatial neglect symptoms in chronic stroke patients improved after 20 sessions of robot-assisted hand training. To confirm the effects of robot-assisted left-hand training on hemispatial neglect in stroke patients, future research should consider randomized controlled trials with various treatment protocols (types of robotic devices and training programs such as passive mode, assistive mode, and active mode).

## Acknowledgment

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) program(IITP-2023-2020-0-01832) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation).

## References

[1] M. Chechlacz, P. Rotshtein, and G.W. Humphreys, Neuroanatomical dissections of unilateral visual neglect symptoms: ALE meta-analysis of lesion-symptom mapping. Frontier in Human Neuroscience 6 (2012), 230.

[2] I. Robertson, P.W. Halligan, C. Bergego, et al, Right neglect following right hemisphere damage? Cortex 3 (1994) 199-213.

[3] V. Varalta, A. Picelli, C. Fonte, et al, Effects of contralesional robot-assisted hand training in patients with unilateral spatial neglect following stroke: a case series study. Journal of Neuroengineering and Rehabilitation 11 (2014), 160.

[4] L. Pizzamiglio, L. Fasotti, M. Jehkonen, et al, The use of optokinetic stimulation in rehabilitation of the hemineglect disorder. Cortex 40 (2004), 441-450.

[5] S. Reinhart, L. Schmidt, C. Kuhn, et al., Limb activation ameliorates body-related deficits in spatial neglect. Frontier in Human Neuroscience 6 (2012), 88.

[6] L.P. Liu, J. Hanly, P. Fahey, et al, A Systematic Review and Meta-Analysis of Rehabilitative Interventions for Unilateral Spatial Neglect and Hemianopia Poststroke From 2006 Through 2016. Archives of Physical Medicine and Rehabilitation 100 (2019), 956-979.

[7] I.H. Robertson and N. North, Spatio-motor cueing in unilateral left neglect: the role of hemispace, hand and motor activation. Neuropsychologia 30 (1992), 553-563.

[8] G. Vallar, M. Rusconi, S. Barozzi, et al, Improvement of left visuo-spatial hemineglect by left-sided transcutaneous electrical stimulation. Neuropsychologia 33 (1995), 73-82.

[9] J.H. Park, Effect of robot-assisted left hand training on unilateral neglect in patients with stroke. Korean Journal of Occupational Therapy 23 (2015), 117-127.

[10] F. Frassinetti, M. Rossi, and E. Ladavas, Passive limb movements improve visual neglect. Neuropsychologia 39 (2001), 725-733.

[11] G.A. Eskes and B. Butler, Using limb movements to improve spatial neglect: the role of functional electrical stimulation. Restorative Neurology and Neuroscience 24 (2006), 385-398.

[12] G. Rizzolatti and A. Berti, Neglect as a neural representation deficit. Revue Neurologique 146 (1990), 626-634.

[13] I.H. Robertson, K. Hogg, and T.M. McMillan, Rehabilitation of unilateral neglect: improving function by contralesional limb activation. Neuropsychological Rehabilitation 8 (1998), 19-29.

[14] J.M. Veerbeek, A.C. Langbroek-Amersfoort, E.E. Van Wegen, et al, Effects of robot-assisted therapy for the upper limb after stroke: a systematic review and meta-analysis. Neurorehabilitation and Neural Repair 31 (2017), 107-121.

[15] P. Sale, M. Franceschini, S. Mazzoleni, et al, Effects of upper limb robot-assisted therapy on motor recovery in subacute stroke patients. Journal of Neuroengineering and Rehabilitation 11 (2014), 104.

[16] R. Maddicks, S.L. Marzillier, and G. Parker, Rehabilitation of unilateral neglect in the acute recovery stage: The efficacy of limb activation therapy. Neuropsychological Rehabilitation 13 (2003), 391-408.

[17] Y.S. Choi, K.W. Lee, J.H. Lee, et al, The effect of an upper limb rehabilitation robot on hemispatial neglect in stroke patients. Annal of Rehabilitation Medicine 40 (2016), 611.

[18] T. Schenkenberg, D. Bradford, and E. Ajax, Line bisection and unilateral visual neglect in patients with neurologic impairment. Neurology 30 (1980), 509-509.

[19] M.L. Albert, A simple test of visual neglect. Neurology 23 (1973), 658-664.

[20] P. Chen, K. Hreha, P. Fortis, et al, Functional assessment of spatial neglect: a review of the Catherine Bergego Scale and an introduction of the Kessler Foundation Neglect Assessment Process. Topic in Stroke Rehabilitation 19 (2012), 423-435.

[21] P. Azouvi, P. Bartolomeo, J-M. Beis, et al, A battery of tests for the quantitative assessment of unilateral. Restorative Neurology and Neuroscience 24 (2006), 273-285.

[22] J. Cohen, Statistical power analysis for the behavioral science. 2nd., Routledge Academic, Hillsdale, MI, 1998.

[23] C.D. Takahashi, L. Der-Yeghiaian, V. Le, and et al, Robot-based hand motor therapy after stroke. Brain 131 (2008), 425-437.

# Identification of Key Variables for Developing a Fall Risk Assessment Tool: Focusing on Statistical Methods and Machine Learning

Chomyong Kim[1], Jung-Yeon Kim[1], Seob Jeon[2], Hyo-Wook Gil[3], Jiwon Lyu[4], Euy Hyun Chung[5], Kwang Seock Kim[6], Yunyoung Nam[7*]

[1]ICT Convergence Research Centre, Soonchunhyang University, Asan, South Korea

[2]Department of Obstetrics and Gynecology, College of Medicine, Soonchunhyang University, Cheonan Hospital, South Korea

[3]Department of Internal Medicine, Soonchunhyang University Cheonan Hospital, South Korea

[4]Division of Respiratory Medicine, Department of Internal Medicine, Soonchunhyang University Cheonan Hospital, South Korea

[5]Department of Dermatology, Soonchunhyang University Cheonan Hospital, South Korea, Soonchunhyang University, Cheonan Hospital, South Korea

[6]Soonchunhyang Innovative Convergence Research Center, Soonchunhyang University, Cheonan Hospital, Asan, South Korea

[7]Department of Computer Science and Engineering, Soonchunhyang University, Asan, South Korea

*Contact: ynam@sch.ac.kr

*Abstract*— **This paper presents a study that identifies critical factors for developing fall risk assessment systems using data from three hospitals and the widely used Morse Fall Scale (MFS). Prediction using machine learning and statistical approaches can be used to identify key variables in developing fall risk assessment tools. The logistic regression and decision tree analyses showed that the decision tree had a high prediction rate of nearly 96%, while logistic regression had an explanatory power of 75.7%. Significant factors for predicting fall risk include behaviour and mobility elements of the pressure ulcer risk assessment tool, drug use, breathing equipment, urinary system medications, daily repositioning, and nutritional intake. These findings can help healthcare facilities in their efforts to prevent falls.**

## I. Introduction

Healthcare is a complex and rapidly evolving industry that requires constant attention to patient safety. Patient falls are a significant concern in healthcare facilities worldwide, as they can lead to severe physical injuries, emotional trauma, and increased healthcare costs [1]. Hospitals and other healthcare organizations have developed various fall risk assessment tools to prevent patient falls. The Morse Fall Scale, the Hendrich II Fall Risk Model, and the St Thomas's Risk Assessment Tool in Falling Elderly Inpatients are among the most widely used fall risk assessment tools [2-4]. Despite their differences, these tools share a common goal of identifying patients at risk of falls and implementing preventive measures. This study aims to identify the critical factors that can be used to develop more effective fall risk assessment tools using the most commonly used tool, the Morse Fall Scale, and electronic medical record (EMR) data. The study will also compare the effectiveness of logistic regression analysis and decision tree analysis using machine learning in identifying fall risk factors. The findings of this study will enhance healthcare facilities' efforts to prevent patient falls and improve patient safety.

## II. Clinical Assessments Tools

In the wards of Seoul, Bucheon, and Cheonan Hospitals affiliated with Soonchunhyang University, various tests, and measures such as MFS-based fall risk assessment, nursing need, pressure sore assessment, pain assessment, and prescription drug records are all recorded on the day of admission.

Excluding the drug prescription record, the patient's condition is evaluated in 2 to 3 stages, and is mainly classified as 'no help needed, some help needed, all help needed'.

This study aims to confirm the importance of variables in developing a fall risk assessment tool by adding the variables in <Table 1> in addition to the MFS fall assessment tool variables through various preceding studies and FGI with specialists.

TABLE I
Clinical Assessment Tools

| Assessment | Name of Assessment |
|---|---|
| Anthropometric survey | Registration number, admission/discharge date. Private corporal, injured corporal, corporal code, height, weight, etc. |
| Nursing Need | Urination and defecation, eating, changing positions, moving out of bed, breathing, nursing, etc. |
| Patient Fall | Fall experience in the past 3 months, gait/mobility, gait assistance, consciousness/mental state, cognitive impairment, heparin lock, etc. |
| Medications | ATC classification code, prescription quantity, number of days, usage, etc. |
| Bed Sore | Nutritional status, level of activity, etc. |

| | |
|---|---|
| Pain | Pain intensity, other than pain area |

## III. Results

The data for this study was collected and pre-processed from three hospitals affiliated with Soonchunhyang University, using 172 fall patients and 400 non-fall patients, with no omission in the evaluation tool results. As only the information of patients without omissions was included, the available data were limited to the information on the day of the patient's fall. Drugs used in hospitals were clustered using the Anatomical Therapeutic Chemical Classification System (ATC) code, and the Charlson Comobity Index (CCI) was used to confirm the severity of the patient's disease.

### *A. Logistic Regression (Statistical Approach)*

Logistic regression is a statistical method used to analyze the relationship between a categorical dependent variable and one or more independent variables, by estimating the probability of occurrence of an event. In regression analysis, β can be used as a weight, while in logistic regression analysis, the odds ratio can be utilized as a weight [5]. The result of the logistic regression is shown in Table 2. Statistically significant variables influencing falls were move out bed, Meal(diet), nutrition, move ability, required specialized treatment, Cardiac medication, opioids, non-narcotic analgesics, sleeping pills and analgesics, and respiratory drugs.

TABLE 2
Result of Logistic Regression

Logit Regression Results

| | | | |
|---|---|---|---|
| Dep. Variable: | 낙상유무 | No. Observations: | 572 |
| Model: | Logit | Df Residuals: | 533 |
| Method: | MLE | Df Model: | 38 |
| Date: | Thu, 10 Nov 2022 | Pseudo R-squ.: | 0.7570 |
| Time: | 12:44:46 | Log-Likelihood: | -84.994 |
| converged: | True | LL-Null: | -349.75 |
| Covariance Type: | nonrobust | LLR p-value: | 7.117e-88 |

| | coef | std err | z | P>\|z\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| const | -1.9284 | 5.947 | -0.324 | 0.746 | -13.585 | 9.728 |
| MFS | 0.0112 | 0.009 | 1.239 | 0.215 | -0.006 | 0.029 |
| total_cci | -0.0796 | 0.097 | -0.816 | 0.414 | -0.271 | 0.111 |
| 키 | -0.0036 | 0.034 | -0.107 | 0.915 | -0.070 | 0.063 |
| 몸무게 | 0.0013 | 0.017 | 0.077 | 0.939 | -0.032 | 0.035 |
| 성별 | -0.5517 | 0.622 | -0.887 | 0.375 | -1.770 | 0.667 |
| 나이 | 0.0044 | 0.018 | 0.245 | 0.807 | -0.031 | 0.040 |
| ADL_배뇨, 배변 | 0.6365 | 0.463 | 1.374 | 0.170 | -0.272 | 1.545 |
| ADL_체위변경 | -0.4849 | 0.627 | -0.773 | 0.439 | -1.714 | 0.744 |
| ADL_침상 밖으로 이동 | 1.4868 | 0.489 | 3.039 | 0.002 | 0.528 | 2.446 |
| ADL_식사섭취 | -1.3631 | 0.508 | -2.683 | 0.007 | -2.359 | -0.367 |
| 호흡 간호_비침습적 산소 투여 | -0.1406 | 0.736 | -0.191 | 0.849 | -1.583 | 1.3 |
| 통증강도 | 0.1605 | 0.127 | 1.269 | 0.205 | -0.087 | 0.408 |
| 욕창_행동 | 1.6486 | 0.387 | 4.260 | 0.000 | 0.890 | 2.407 |
| 욕창_영양 | 0.9894 | 0.783 | 1.263 | 0.206 | -0.545 | 2.524 |
| 욕창_기동력 | 2.2573 | 0.502 | 4.497 | 0.000 | 1.274 | 3.241 |
| ADL_위험 | 2.7757 | 1.604 | 1.731 | 0.083 | -0.368 | 5.919 |
| ADL_전문 | 1.4063 | 0.681 | 2.065 | 0.039 | 0.072 | 2.741 |
| Cluster_A | -0.9504 | 0.472 | -2.014 | 0.044 | -1.875 | -0.026 |
| Cluster_B | -0.1103 | 0.350 | -0.315 | 0.753 | -0.797 | 0.576 |
| Cluster_C | -0.6901 | 0.275 | -2.506 | 0.012 | -1.230 | -0.150 |
| Cluster_D | -0.8801 | 0.428 | -2.054 | 0.040 | -1.720 | -0.040 |
| Cluster_E | 0.1589 | 0.251 | 0.632 | 0.528 | -0.334 | 0.652 |
| Cluster_F | -0.2573 | 0.390 | -0.659 | 0.510 | -1.022 | 0.508 |
| Cluster_G | -0.1356 | 0.238 | -0.571 | 0.568 | -0.601 | 0.330 |
| Cluster_H | -0.4313 | 0.394 | -1.094 | 0.274 | -1.204 | 0.342 |
| Cluster_I | -0.1989 | 0.370 | -0.537 | 0.591 | -0.924 | 0.527 |
| Cluster_J | -0.1777 | 0.493 | -0.361 | 0.718 | -1.143 | 0.788 |
| Cluster_K | 0.7871 | 0.556 | 1.415 | 0.157 | -0.303 | 1.878 |
| Cluster_L | -0.2265 | 0.361 | -0.627 | 0.531 | -0.935 | 0.482 |
| Cluster_M | 0.0102 | 0.492 | 0.021 | 0.984 | -0.955 | 0.975 |
| Cluster_N | 0.2574 | 0.538 | 0.479 | 0.632 | -0.796 | 1.311 |
| Cluster_O | -0.7114 | 0.707 | -1.006 | 0.314 | -2.097 | 0.674 |
| Cluster_P | 1.6227 | 0.763 | 2.126 | 0.033 | 0.127 | 3.119 |
| Cluster_Q | 1.0391 | 0.700 | 1.485 | 0.138 | -0.333 | 2.411 |
| Cluster_R | -0.0116 | 0.316 | -0.037 | 0.971 | -0.631 | 0.607 |
| Cluster_S | -1.2829 | 0.423 | -3.033 | 0.002 | -2.112 | -0.454 |
| Cluster_T | 0.2667 | 0.431 | 0.620 | 0.536 | -0.577 | 1.111 |
| Cluster_U | -1.1777 | 0.921 | -1.279 | 0.201 | -2.983 | 0.627 |

To standardize, the mean of β is computed for each variable, and then the weight of each variable is derived by dividing its β value by the mean. The risk of falling can be evaluated by adding the constant and the weight of each variable, as shown in Table 3. Although the beta values were not significantly different, patients with dementia and those experiencing flash symptoms, as well as those taking sleeping pills and sedatives, were found to have the highest weight, which aligns with the nurses' professional opinions. The only variable that was statistically significant and had a high weight was Move Ability.

TABLE 3
Calculation with Weights from Logistic Regression

| | Variables | Weight | | Variables | Weight |
|---|---|---|---|---|---|
| | Const | 0.056805 | 20 | Opioids* | 0.195963 |
| 1 | MFS | 0.395135 | 21 | Other analgesics* | 0.162054 |
| 2 | CCI | 0.360838 | 22 | Constipation, antidiarrheals | 0.458027 |
| 3 | Height | 0.38933 | 23 | Steroids | 0.302091 |
| 4 | Weight | 0.391243 | 24 | Antithrombotic agents | 0.341186 |
| 5 | Gender | 0.225051 | 25 | Antihistamines | 0.253846 |
| 6 | Age | 0.392457 | 26 | Antihypertensives | 0.320258 |
| 7 | Urination/Defecation | 0.738431 | 27 | Anxiolytics | 0.32712 |
| 8 | Position Change | 0.240598 | 28 | Antiepileptics | 0.85845 |
| 9 | Move out of bed** | 1.728187 | 29 | Lipid modifying agents | 0.31154 |
| 10 | Meal** | 0.099976 | 30 | Blood glucose lowering | 0.39474 |
| 11 | Respiratory Care | 0.339484 | 31 | Antidepressants | 0.505439 |
| 12 | Pain Intensity | 0.45876 | 32 | Anti-psychotics | 0.191834 |
| 13 | Activity | 2.0317 | 33 | Hypnotics & sedatives* | 1.979754 |
| 14 | Nutrition | 1.050927 | 34 | Muscle relaxants | 1.104478 |
| 15 | Move Ability*** | 3.734346 | 35 | Urological | 0.386228 |
| 16 | Dementia, Delirium | 6.271231 | 36 | Respiratory drugs** | 0.108324 |
| 17 | Professional Treatment* | 1.59452 | 37 | Nonsteroidal anti-inflammatory | 0.510162 |
| 18 | Cardiac drugs* | 0.151053 | | Antineoplastic agents | 0.120341 |
| 19 | Diuretics | 0.349928 | | | |

### *B. Decision Tree (Machine Learning)*

The decision tree model is a type of analysis that predicts output values by applying a series of predictable rules to patterns within the data. This model is known as Classification And Regression Tree (CART) as it is capable of both classification and regression analysis.

In this study, the model was optimized through hyperparameter tuning with the 'max_depth' set to 5, 10, 15, and 20, and 'min_samples_leaf' set to 5, 10, and 15. The resulting cross-analysis demonstrated a predictive power of 93%.

Decision trees quantify the contribution of each input variable or feature to the model's output by referring to this as feature significance. It evaluates the importance of each

characteristic in forecasting the target variable and aids in locating the features that have the greatest influence on the accuracy of predictions. The reduction in impurity, such as the Gini impurity or entropy, that follows from dividing a node based on a specific characteristic is often used to quantify the relevance of a feature. The trait is regarded as being more significant the higher the impurity decrease. To choose the most pertinent characteristics for creating a predictive model and to learn more about the underlying connections between the input and output variables, decision trees' feature importance can be employed [6]. In contrast to the findings of Logistic Regression, the feature importance of the Decision Tree revealed that CCI, weight, age, diarrhea and laxatives, antiplatelet drugs, urinary system drug intake, and respiratory system drug intake were significant variables in model prediction, while other variables like move ability and activity, as well as narcotics, were frequently observed.

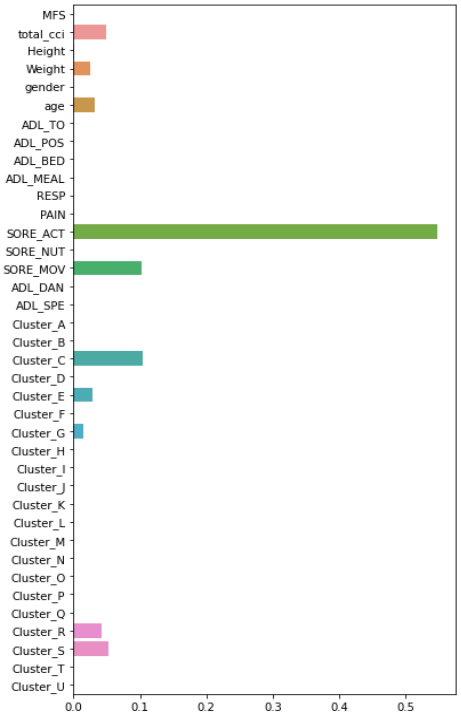



Fig. 1 Feature Importance of Decision Tree Classifier

## IV. Conclusions

In order to choose the factors needed for creating a fall risk assessment tool, this study from Soonchunhyang University used electronic medical data and several assessment tools from three hospitals. For variable selection, machine learning techniques such as logistic regression and decision trees were used. The explanatory power of logistic regression was 75.7%, which is comparable to the current accuracy of Soonchunhyang University's MFS. The observed result can be attributed to two factors. Firstly, the data ratio used in the study was randomly selected and may not have been suitable for the model due to the use of imbalanced data. Secondly, the study did not consider the temporal effects of drugs. The data used in the study were collected on the day of the fall, but the effects of drugs can last for up to three days, and thus, expanding the temporal scope of the study would be reasonable. Logistic regression detected statistically significant factors by p-value and extracted variables corresponding with nurses' experience through regression coefficient calculation. In comparison, Decision Tree attained a predictive power of up to 93%, however when other machine learning approaches are utilized, the variable importance can vary, making it challenging to prove statistical significance. Despite the challenges posed by explainable artificial intelligence, it is important not to overlook the potential benefits of AI analysis and its findings. Nevertheless, it is necessary to conduct thorough studies, publish results in peer-reviewed journals, and validate the clinical implications in real-world settings before implementing and applying them[7].

## Acknowledgment

This study received approval from the Institutional Review Board (IRB number: SCHCA2021-12-029)

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) program(IITP-2023-2020-0-01832) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation)

## References

[1] Najafpour, Z., Godarzi, Z., Arab, M., & Yaseri, M. (2019). Risk Factors for Falls in Hospital In-Patients: A Prospective Nested Case Control Study. *International journal of health policy and management*, *8*(5), 300–306. https://doi.org/10.15171/ijhpm.2019.11

[2] Schwendimann, R., De Geest, S., & Milisen, K. (2006). Evaluation of the Morse Fall Scale in hospitalised patients. Age and ageing, 35(3), 311-313.

[3] Hendrich, A. L., Bender, P. S., & Nyhuis, A. (2003). Validation of the Hendrich II Fall Risk Model: a large concurrent case/control study of hospitalized patients. Applied Nursing Research, 16(1), 9-21.

[4] Oliver, D., Britton, M., Seed, P., Martin, F. C., & Hopper, A. H. (1997). Development and evaluation of evidence-based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: case-control and cohort studies. Bmj, 315(7115), 1049-1053.

[5] Lee, S. J., Kim, Y. T., & Kim, S. Y. (2013). Comparison of Customer Satisfaction Indices Using Different Methods of Weight Calculation. Journal of Digital Convergence, 11(12), 201-211.

[6] Saarela, M., & Jauhiainen, S. (2021). Comparison of feature importance measures as explanations for classification models. *SN Applied Sciences, 3*, 1-12.

[7] Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

# Hypotension prediction from invasive blood pressure using vision transformer wavelet image

Geon Lee[1,*], JiYoung Woo[2]
[1]*Dept. ICT convergence, University of Soonchunhyang, Asan, South Korea*
[2]*Dept. AI and Bigdata, University of Soonchunhyang, Asan, South Korea*
*Contact: jywoo@sch.ac.kr, phone +1-778 782 0000

*Abstract*— **Hypotension can result in serious complications if not promptly detected and managed. In this study, we present a new approach for predicting hypotension using invasive blood pressure (IBP) signals, employing wavelet transform and vision transformers. Wavelet transform is utilized to convert IBP signals into time-frequency domain images, capturing both the temporal and spectral features of the signals. In this study, we use 4 types of wavelet transform image. These images are used as input for a vision transformer. Vision transformer, using the transformer architecture initially developed for natural language processing, handles images by dividing them into fixed-size patches and linearly embedding them into a sequence of tokens. The transformer layers capture complex patterns and relationships among these image patches, enabling effective and efficient image classification and analysis. This research demonstrates that it's possible to predict hypotension from invasive blood pressure signals not by extracting thousands of variables, but simply through the process of converting these signals into images. This emphasizes the potential of vision transformers in the medical field and may contribute to improved patient monitoring and clinical decision-making in critical care environments.**

## I. INTRODUCTION

Invasive and non-invasive blood pressure measurements are two prevalent methods for assessing blood pressure. Invasive blood pressure measurement, which requires the insertion of a catheter into the body, is commonly used during surgical procedures. Monitoring blood pressure during surgery is crucial, as fluctuations can significantly impact the patient's health. For example, hypotension during surgery may elevate the risk of heart or kidney damage, with the associated risks increasing as the duration of hypotension extends [1]. Consequently, predicting hypotension during surgery is essential for ensuring patient safety.

One of the most effective algorithms for predicting hypotension to date involves extracting features from arterial pressure waveforms and using machine learning to estimate the likelihood of hypotension occurring within the next 5, 10, or 15 minutes. In this approach, hypotension is defined as a mean arterial pressure (MAP) of less than 65 mmHg. This research has been commercialized as HPI and is currently being applied for real-time hypotension prediction in intensive care units and operating rooms [2].

The process of extracting features from signals often requires considerable expertise and effort. In recent years, there has been a growing interest in using wavelet images as direct inputs for convolutional neural network (CNN) models [3] or vision transformers across various signal processing fields. Although the performance of these novel approaches has not yet surpassed that of traditional feature extraction techniques, their innovation as a new approach holds potential significance.

In the present study, our objective is to apply this state-of-the-art approach to invasive blood pressure monitoring by converting blood pressure data into wavelet images and subsequently training a vision transformer model. Moreover, to enhance the model's predictive capabilities, we incorporate several distinct mother wavelets simultaneously. Our predictive model is capable of anticipating hypotensive events one minute in advance.

## II. EXPERIMENTS

### A. Dataset

In this study, the data utilized were collected from the operating rooms of Soonchunhyang University Hospital. The data was acquired using the GE Bx50 monitor and includes invasive blood pressure, electrocardiogram (ECG), plethysmograph (PLETH), and carbon dioxide (CO2) measurements. A total of 7,321 data points were gathered at a sampling rate of 100 Hz from 16 patients.

### B. Data Preprocessing

IBP data is segmented using a sliding window of 140 seconds in length with 60 seconds lookback. From the segmented data, the first 20 seconds are used for hypotension prediction, while the following 60 seconds are utilized for hypotension labeling. Segmented data with more than 10% missing values, no variation at specific values, or exhibiting excessively large fluctuations are considered outliers and removed from the analysis. After preprocessing, data points with a MAP of 65 mmHg or lower are labeled as hypotensive, while those with higher values are labeled as normal.

### C. Wavelet Transform

Wavelet transformation is a technique that enables the analysis of the frequency domain with respect to time variations. Wavelet transform involves the use of a mother wavelet function to calculate the correlation coefficients between the original signal and the function while moving along the time axis. The process is repeated by adjusting the scale of the mother wavelet function and moving it again in the time direction to compute the correlation coefficients. The choice of mother wavelet can yield diverse analysis results,

so selecting an appropriate mother wavelet is crucial. One distinguishing aspect of this study is the application of various mother wavelet to extract different features, which are then analyzed as images. This approach allows for a more comprehensive examination of the distinct characteristics associated with each mother wavelet.

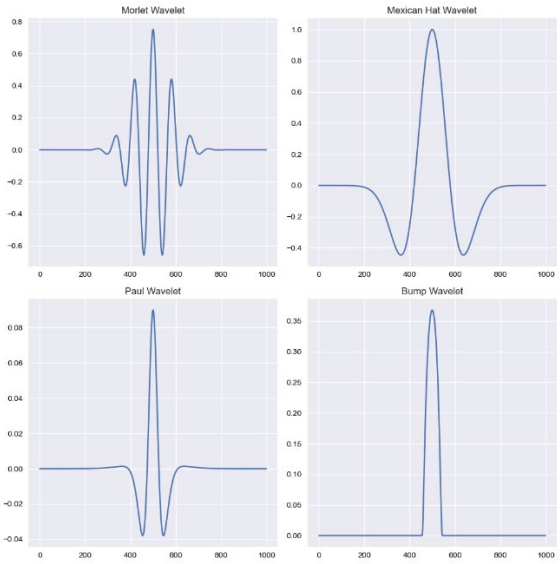



Fig. 1 Four types of mother wavelet (Morlet, Mexican, Paul, Bump)

Fig. 1 shows four types of mother wavelet utilized in this study. According to existing research findings, it has been demonstrated that Bump and Morlet mother wavelets correspond more closely with theoretical predictions, whereas the Mexican hat effectively captures short-period variability, and the Paul wavelet exhibits long-period variability [4]. As a result, we utilized wavelet transform images generated by each of these mother wavelets to extract unique signal characteristics.

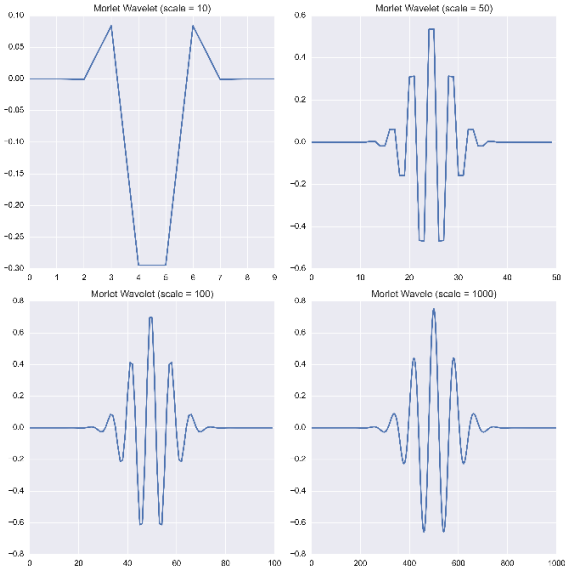



Fig. 2 Morlet wavelets with different scales

Fig. 2 shows Morlet wavelet with different scales. There is a noticeable difference when the scale is 10, 50, and 100, but the wavelets appear similar when the scale is 100 and 1000. However, in this case, the scale size of the x-axis differs; assuming the original signal is at 100 Hz, the mother wavelet is applied at 1-second intervals for a scale of 100 and at 10-second intervals for a scale of 1000.

Thus, the objective of this study is to extract various features from invasive blood pressure signals using a diverse range of scales and mother wavelets.

### *D. Research Design*

First, we perform a wavelet transform on the signal data consisting of 2,000 data points over 20 seconds at 100 Hz. We employ four types of mother wavelets to generate four different wavelet transform images, adding dimensions to learn these images simultaneously. Subsequently, the data is restructured from its original 100x2000-sized image format into a 100x100-sized image. This process is carried out to improve computational efficiency and facilitate the input of fixed-size patches into the Vision Transformer [5].

The vision transformer is an innovative architecture for image processing, applying the Transformer, which was originally utilized in the natural language processing domain, to the field of image processing. Initially, the entire image is divided into small patches of a fixed size and flattened into a vector format using image patches. Following this, position embeddings are added to each patch to incorporate location information. These combined vectors then pass through the Transformer encoder layer to extract the features of the image.

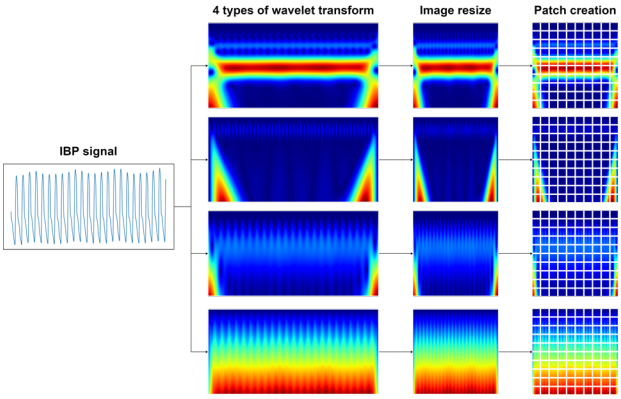



Fig. 3 Signal to patch process

Fig. 3 shows the procedure of segmenting the 100x100 image into 10x10-sized patches. These subdivided images are then flattened and combined with position embeddings before being processed through the Transformer encoder. This step is depicted in Fig. 4.

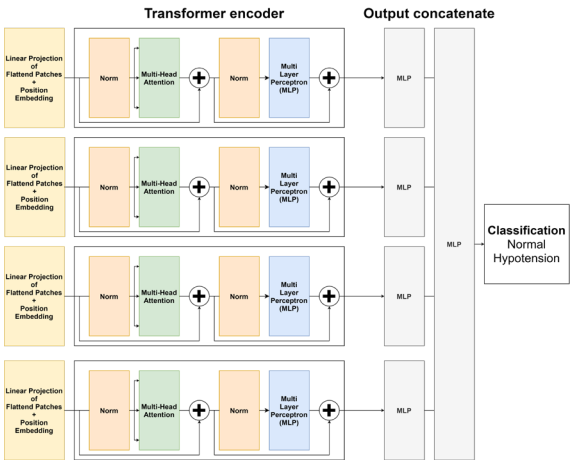



Fig. 3 Transformer process

In this study, after applying each of the images to the Vision Transformer, the resulting vectors are concatenated. The concatenated data is then processed through a multilayer perceptron (MLP) layer, ultimately utilized for the task of classifying Normal and Hypotension cases.

## III. RESULTS

In this study, we compare our findings with the metrics of a related research paper that also utilizes invasive blood pressure data [6]. In the referenced study, a machine learning model was developed by applying various signal processing techniques, including general statistics, peak analysis, Fourier analysis, and wavelet analysis.

We compared our vision transformer model with two other models: one using all features and another utilizing Peak, Fourier, and Wavelet features in the related research paper. The comparison details can be found in TABLE 1.

TABLE I
COMPARE RELATED RESEARCH AND OUR RESULTS

| method | Precision | Recall | F1 score |
|---|---|---|---|
| Statistics + Peak + Fourier + Wavelet | 0.652 | 0.441 | 0.526 |
| Peak + Fourier + Wavelet | 0.656 | 0.305 | 0.416 |
| Early_stopping = 30, Learning_rate = 0.001 Undersampling | 0.297 | 0.961 | 0.454 |
| Early_stopping = 50, Learning_rate = 0.002 Undersampling | 0.635 | 0.414 | 0.501 |
| Early_stopping = 70, Learning_rate = 0.002 Undersampling | 0.419 | 0.991 | 0.574 |

The values above the double line in TABLE 1 represent evaluation metrics from other studies that focus on the same data. The values below the double line indicate the evaluation metrics for the approach employed in this research.

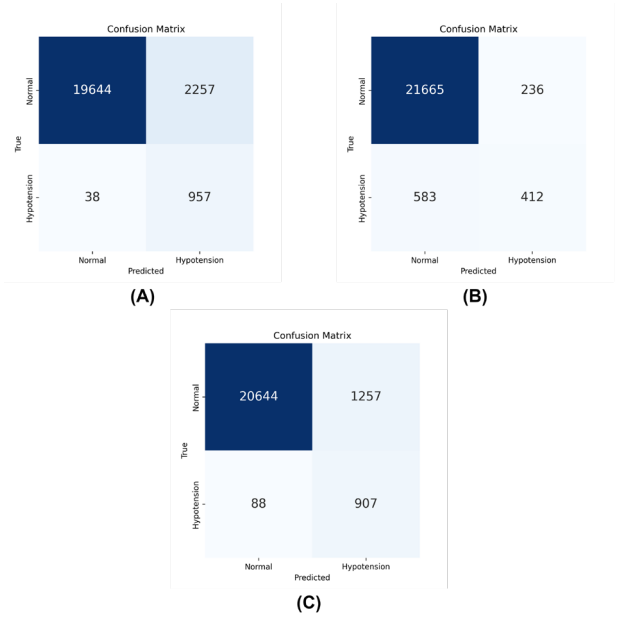



Fig. 4 Confusion matrix

In Fig. 4, the confusion matrix represents the results of different models: (A) illustrates the model with early stopping set to 30 and learning rate at 0.001, (B) demonstrates the model with early stopping at 50 and learning rate of 0.002, and finally, (C) showcases the model with early stopping set to 70 and learning rate at 0.002. Because of the significant class imbalance in the data, undersampling was applied to all models.

All models exhibited a tendency to misclassify a large number of normal cases as hypotension in order to correctly predict a higher proportion of hypotension cases. In particular, model (A) had a severely imbalanced ratio, resulting in a recall of 0.961 but a precision of only 0.297. While model (B) reduced overall misclassifications, it struggled to accurately predict hypotension cases. Model (C) demonstrated a similar trend to model (A), but with a significant decrease in the number of normal instances misclassified as hypotension.

## IV. CONCLUSIONS

In this study, we conducted an experiment to predict hypotension by applying various wavelet transform images to the Vision Transformer. While the performance of the current study did not surpass that of previous research, it was found that classification could still be achieved to some extent using only wavelet transform without the need for additional signal characteristics. In future work, combining wavelet transform with other blood pressure data features or fully utilizing the data removed due to imbalance could potentially provide a novel approach to predicting hypotension.

## ACKNOWLEDGEMENT

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2023-2020-0-01832) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation).

## REFERENCES

[1] Salmasi et al., Relationship between Intraoperative Hypotension, Defined by Either Reduction from Baseline or Absolute Thresholds, and Acute Kidney and Myocardial Injury after Noncardiac Surgery. Anesthesiology 2017.

[2] Hatib et al., Machine-learning Algorithm to Predict Hypotension Based on High-fidelity Arterial Pressure Waveform Analysis. Anesthesiology 2018.

[3] Keerthi Krishnan, K., Soman, K.P. CNN based classification of motor imaginary using variational mode decomposed EEG-spectrum image. Biomed. Eng. Lett. 11, 235–247 (2021).

[4] Lee, Jinwook, Hyunwook Lee, and Chulsang Yoo. "Selection of mother wavelet for bivariate wavelet analysis." Journal of Korea Water Resources Association 52.11 (2019): 905-916.

[5] Dosovitskiy, Alexey, et al. "An image is worth 16x16 words: Transformers for image recognition at scale." arXiv preprint arXiv:2010.11929 (2020).

[6] Subin Lee, Hyung-Chul Lee, SangHyun Kim, and JiYoung Woo. "Development of a Prediction Model for Hypotension Generation in Surgical Patients Using Invasive Blood Pressure Data" Proceedings of the Korean Information Science Society Conference, (2021) : 1031-1032.

# Multi-channel EEG-based drowsy driving state classification using InceptionTime

Minsu Cha[1], Jiyoung Woo[1]
[1]*University of Soonchunhyang, Dept. ICT convergence, Asan, South Korea*
*Contact: ckalstn0522@gmail.com, jywoo@sch.ac.kr

*Abstract*— **Drowsy driving has become a significant issue in traffic accidents in recent years. However, traditional drowsy driving prediction studies have primarily relied on single-channel electroencephalogram (EEG) data to classify drivers' drowsiness states. This study proposes a drowsy driving prediction method that utilizes multi-channel EEG data and the InceptionTime model, a deep learning architecture designed for time series classification. This method provides an advantage over single-channel models by employing multi-channel data, leading to more accurate drowsiness state classification and improved drowsy driving prevention. Additionally, the InceptionTime model can capture various frequency and time-domain patterns of multi-channel EEG data, thus enabling higher performance than previous methods. This study aims to demonstrate the effectiveness of multi-channel EEG data for drowsy driving prediction and confirm the high performance achieved by applying the InceptionTime model.**

## I. Introduction

Drowsy driving has emerged as a critical issue in road safety, contributing to many traffic accidents worldwide. In recent years, the increasing prevalence of drowsy driving has drawn attention to the need for developing effective methods to predict and prevent such incidents. According to the American Automobile Association (AAA), over 40% of drivers have reported experiencing drowsiness while driving, which highlights the severity of this issue. Moreover, a study conducted by the AAA discovered that drowsy driving contributes to 9.5% of all vehicle accidents and 10.8% of accidents involving airbag deployment, injuries, or significant property damage. The National Safety Council (NSC) reports that drowsy driving leads to an estimated 100,000 vehicle accidents annually, resulting in approximately 71,000 injuries and 1,550 fatalities. According to a survey by the Centers for Disease Control and Prevention (CDC), one in 25 drivers admitted to experiencing drowsiness while driving. The National Highway Traffic Safety Administration (NHTSA) reported that in 2020 alone, there were 633 fatalities due to drowsy driving-related traffic accidents.[1]-[3]

Various studies have explored drowsiness detection using different methodologies. V. P. B. Venkat et al.[4], [5] employed wavelet packet transform (WPT) in a single-channel EEG-based drowsiness detection model and also introduced a novel feature extraction strategy based on a single Hjorth parameter, surpassing existing Power Spectral Density (PSD) features. Jian Cui et al.[6] developed a Convolutional Neural Networks-Long Short Term Memory(CNN-LSTM) model for subject-independent drowsiness recognition from single-channel EEG signals. Ali Ghadami et al.[7] proposed a 1D-CNN model for automatic drowsiness detection using single-channel EEG spectrograms. H. Albalawi et al.[8] applied a Support Vector Machine (SVM) in a real-time drowsiness detection algorithm based on single-channel EEG for wearable devices, achieving superior accuracy over the conventional method.

While these previous studies focused on single-channel approaches, our research expands on these methodologies by utilizing multi-channel EEG data, offering additional advantages in accuracy and robustness for drowsiness detection. Furthermore, our study investigates optimal multi-channel combinations that can deliver superior performance in detecting drowsy driving, ultimately enhancing the effectiveness and reliability of the drowsiness detection system.

In this research, we aim to address the constraints of previous drowsy-driving prediction methods by proposing a novel approach using multi-channel EEG data. Specifically, we employ the InceptionTime model, a deep learning architecture designed for time series classification, to analyze and classify drowsiness states based on multi-channel EEG data. Using multi-channel data provides an advantage over single-channel models, leading to more accurate drowsiness state classification and improved drowsy driving prevention. Moreover, the InceptionTime model can capture various frequency and time-domain patterns of multi-channel EEG data, thus enabling higher performance than previous methods.[9]

## II. Materials and Methods

In this study, we utilized an EEG dataset published by Cao et al.[10], which was collected from 27 participants who engaged in a virtual reality driving simulation.

The EEG dataset was acquired using 30 channels with a sampling rate of 500 Hz. Artifacts were processed to ensure the data's quality and reliability.

Subsequently, the dataset was downsampled to 128 Hz.[11] Each extracted sample consisted of 3 seconds of EEG data.

During the simulation, lane departure events occurred. The time taken for participants to steer their vehicle back to the center of the lane was defined as the 'reaction time'. The participants' EEG data were labeled as 'alertness' or 'drowsiness' based on their reaction times. In our analysis, we used data from 11 of the 27 participants to ensure a balanced

representation of 'alertness' and 'drowsiness' labels. We chose these 11 participants because their data exhibited the most balanced distribution of these labels.

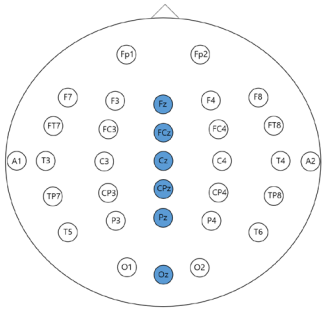



Fig. 1 Electrodes were positioned following the international 10-20 system. F: denotes the frontal lobe, T: represents the temporal lobe, C: signifies the central lobe, P: indicates the parietal lobe, and O: corresponds to the occipital lobe. Z refers to an electrode positioned along the midline. [12]

TABLE I
Performance comparison table among single-channel, O-Group (O1, Oz, O2), and Z-Group (Fz, FCz, Cz, CPz, Pz, Oz) channels.

| Subjects | Single (Oz) | O-Group | **Z-Group** |
|---|---|---|---|
| S1 | 73.93 | 84.57 | 83.51 |
| S2 | 51.51 | 56.81 | 73.48 |
| S3 | 58.66 | 58.66 | 72.66 |
| S4 | 68.91 | 70.94 | 69.59 |
| S5 | 65.62 | 58.92 | 83.48 |
| S6 | 83.73 | 81.92 | 75.30 |
| S7 | 63.72 | 64.70 | 72.54 |
| S8 | 59.46 | 59.84 | 68.93 |
| S9 | 79.29 | 73.56 | 87.26 |
| S10 | 83.33 | 87.03 | 87.03 |
| S11 | 72.12 | 70.35 | 71.23 |
| Average | 69.12 | 69.76 | **76.82** |

Our study selected six EEG channels placed along the Z-line according to the international 10-20 system. The chosen channels were Fz, FCz, Cz, CPz, Pz, and Oz. These channels were selected based on the highest accuracy achieved in our experiments, where we grouped channels by their location according to the symbols shown in Fig. 1 and tested their performance in classifying drowsiness states. In our analysis, we examined the performance of three groups: two groups that included the Oz channel, which was used for labeling, and another group that used single-channel data.

Following this, the Short-Time Fourier Transform (STFT) was utilized to analyze the frequency components of the EEG data while preserving the time information. The STFT was applied to the dataset with a sampling frequency of 384 Hz and a segment length of 64, enabling a detailed examination of the frequency-domain patterns in the EEG data.

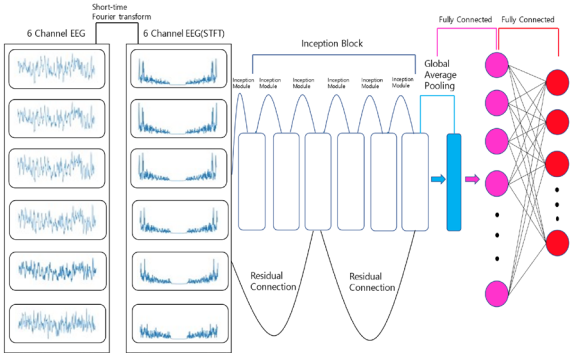



Fig. 2 Data and InceptionTime model architecture

Utilizing the InceptionTime model on the multi-channel EEG data provides several advantages. The bottleneck layer of the InceptionTime model significantly reduces the dimensionality of the time series, mitigating the model's complexity and addressing overfitting problems in small datasets. By simultaneously applying multiple filters of different lengths on the same input time series, the InceptionTime model can capture complex patterns and relationships in multi-channel data.[9]

These features of the InceptionTime model enable it to extract multi-resolution characteristics from multi-channel EEG data, improving accuracy in drowsiness state classification and enhancing drowsy driving prevention. This performance enhancement is notable compared to previous methods that relied on single-channel data, establishing the InceptionTime model as a powerful tool for analyzing multi-channel data.

To assess the effectiveness of our approach in comparison to prior studies, we performed a leave-one-out cross-validation. The model was trained for 8 epochs with a batch size of 32, allowing us to evaluate its performance.

## III. Results

In the results section, we compared the accuracy of the proposed model with traditional methods such as Linear Discriminant Analysis(LDA), Quadratic Discriminant Analysis(QDA), SVM, as well as a CNN model from previous research[11], in terms of drowsiness state classification accuracy.

TABLE IIIII
Performance comparison table between the model proposed in this study and the models used in previous research.

| Subjects | LDA | QDA | SVM | CNN | **Proposed Model** |
|---|---|---|---|---|---|
| S1 | 75.00 | 77.66 | 77.13 | 77.45 | **83.51** |
| S2 | 43.94 | 39.39 | 46.21 | 52.80 | **73.48** |
| S3 | 49.33 | 51.33 | 49.33 | 63.47 | **72.66** |
| S4 | 50.68 | 58.11 | 61.49 | **76.22** | 69.59 |
| S5 | 64.29 | 60.27 | 68.75 | 76.52 | **83.48** |
| S6 | **87.95** | 83.13 | 85.54 | 77.11 | 75.30 |
| S7 | 65.69 | 64.71 | 63.73 | 67.35 | **72.54** |
| S8 | **77.65** | 71.97 | 73.48 | 71.93 | 68.93 |
| S9 | 76.11 | 78.34 | 81.21 | **88.25** | 87.26 |
| S10 | 87.96 | **87.96** | 86.11 | 81.67 | 87.03 |
| S11 | 65.04 | 65.04 | 69.03 | **72.56** | 71.23 |
| Average | 67.60 | 67.08 | 69.27 | 73.22 | **76.82** |

Table 1 showcases the performance of the traditional methods, the CNN model, and the model proposed in our study. Although our model achieves the highest individual accuracy in only five cases, it demonstrates the highest average accuracy of 76.82 while maintaining strong generalization performance. This indicates that the proposed model outperforms the traditional methods and the previous research's CNN model regarding accuracy and generalization.

## IV. Conclusions

In conclusion, our study successfully addresses the limitations of previous drowsy-driving prediction methods by employing a novel approach using multi-channel EEG data and the InceptionTime model. This deep learning architecture for time series classification allows for more accurate drowsiness state classification and improved drowsy driving prevention. By focusing on six EEG channels along the Z-line according to the international 10-20 system, we achieved higher accuracy and better performance in our classification task than using other channels or a larger number of channels.

The results of our study indicate that the proposed model surpasses traditional methods and the CNN model from previous research in terms of accuracy and generalization. With an average accuracy of 76.82, our model exhibits strong generalization performance, indicating its potential for practical application in real-world scenarios.

Future research in this area could further explore integrating additional data sources, such as physiological and behavioral data, to enhance the robustness and accuracy of drowsiness detection systems. Moreover, developing real-time drowsiness monitoring systems that can be easily implemented in vehicles could significantly reduce drowsy driving-related accidents and improve overall road safety.

## ACKNOWLEDGEMENT

This research was supported by “Regional Innovation Strategy (RIS)” through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE)(2021RIS-004).

## REFERENCES

[1] (2022) Drowsy driving statistics and facts 2022 | Bankrate [Online]. Available: https://www.bankrate.com/insurance/car/drowsy-driving-statistics/

[2] (2020) Drowsy Driving: Avoid Falling Asleep Behind the Wheel | NHTSA [Online]. Available: https://www.nhtsa.gov/risky-driving/drowsy-driving

[3] (2020) Sleep and Sleep Disorders | CDC [Online]. Available: https://www.cdc.gov/sleep/features/drowsy-driving.html

[4] V. Phanikrishna B and S. Chinara, "*Automatic classification methods for detecting drowsiness using wavelet packet transform extracted time-domain features from single-channel EEG signal*," Journal of Neuroscience Methods, vol. 347, Jan. 2021.

[5] V. Phanikrishna B and S. Chinara, "Time Domain Parameters as a feature for single-channel EEG-based drowsiness detection method," 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 2020, pp. 1-5

[6] J. Cui et al., "Subject-Independent Drowsiness Recognition from Single-Channel EEG with an Interpretable CNN-LSTM model," 2021 International Conference on Cyberworlds (CW), Caen, France, 2021, pp. 201-208

[7] A. Ghadami, M. Mohammadzadeh, M. Taghimohammadi, and A. Taheri*, "Automated Driver Drowsiness Detection from Single-Channel EEG Signals Using Convolutional Neural Networks and Transfer Learning,"* 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), Macau, China, 2022, pp. 4068-4073

[8] H. Albalawi and X. Li*, "Single-Channel Real-Time Drowsiness Detection Based on Electroencephalography,"* 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 2018, pp. 98-101

[9] H.I. Fawaz, B. Lucas, G. Forestier, C. Pelletier, D.F. Schmidt, J. Weber, G.I. Webb, L. Idoumghar, P.A. Muller, F. Petitjean, *“Inceptiontime: Finding alexnet for time series classification,”* Data Mining and Knowledge Discovery 34(6), 1936–1962, 2020

[10] Cao, Z., Chuang, CH., King, JK. et al. *“Multi-channel EEG recordings during a sustained-attention driving task,”* Sci Data 6, 19, 2019

[11] J. Cui et al. “A Compact and Interpretable Convolutional Neural Network for Cross-Subject Driver Drowsiness Detection from Single-Channel EEG,” Methods, Volume 202, Pages 173-184, 2022

[12] Pal, N.R., et al., “EEG-based subject-and session-independent drowsiness detection: an unsupervised approach,” EURASIP Journal on Advances in Signal Processing, 2008. 2008(1): p. 519480.

# Real-time Prediction of Anaesthesia Complications Using a Transformer on Non-Invasive Vital Sign

EunSeo Jung[1,*], JiYoung Woo[2], and SangHyun Kim[3]
[1]*Dept. ICT Convergence, University of Soonchunhyang, Asan, South Korea*
[2]*Dept. AI and Bigdata, University of Soonchunhyang, Asan, South Korea*
[3]*Soonchunhyang University Bucheon Hospital, Bucheon, South Korea*
*Contact: jywoo@sch.ac.kr

*Abstract*— **The study aims to develop a real-time prediction model for hypotension and hypertension in patients during anaesthesia for practical use in clinical settings. We collected data from 516 patients and selected variables based on their correlation coefficients with blood pressure and surgical stages of Induction, Intubation and Operation. Our contribution is developing a model that can predict blood pressure-related events 5 minutes in advance using only non-invasive intraoperative vital sign data and intravenous anaesthesia information without needing any prior patient information. We proposed a Transformer model that was trained to classify patients' status into three categories: hypotension, hypertension, and normal blood pressure. The model achieved an overall accuracy of 81% and an f1-score of 62%. The results indicate that the proposed Transformer model can effectively predict blood pressure and enable medical personnel to take preventive measures towards patients.**

## I. Introduction

Fluctuations in blood pressure during general anaesthesia can adversely affect a patient's prognosis, and blood pressure-related events can lead to complications such as myocardial injury. Therefore, detecting anaesthesia complications in real-time and providing early warning could enable medical personnel to take proactive measures, prevent adverse effects and potentially rescue patients' lives [1].

In a prior study [2], researchers developed CNN (Convolutional Neural Network) models to forecast hypotensive events in patients 1 minute in advance. Furthermore, in a separate study [3], an LSTM (Long Short-Term Memory) model was employed to predict hypotensive events 5 minutes before their occurrence in patients undergoing colorectal surgery. The model utilized intraoperative vital sign data and patient information, including age and pre-existing hypertension, to enhance the accuracy of the predictions. In a previous study on non-invasive blood pressure prediction, as cited in [4], vital signs such as PPG (Photoplethysmography) and ECG (Electrocardiography) were utilized to predict blood pressure values.

This study aims to provide a practical solution for the urgent need to predict and prevent anaesthesia complications in real-time. The existing literature on predicting blood pressure changes during anaesthesia has relied on predicting such changes 1 minute before they occur or requiring additional patient data such as age and medical history. However, this study proposes a novel approach that utilizes only intraoperative data, which can be easily collected during surgery, to predict blood pressure changes up to 5 minutes in advance. This model offers a promising approach to enhancing patient safety during general anaesthesia, potentially reducing the incidence of anaesthesia-related complications and improving overall outcomes.

## II. Experiments

### A. Data

We collected datasets from 736 patients undergoing general anaesthesia at Soonchunhyang University Bucheon Hospital in Korea from March 2, 2021, to June 30, 2021 (approval No. SCHDB_IRB_2011-11-015). However, we excluded datasets from patients without recorded surgical stages and those with surgical durations less than 30 minutes because anaesthesia takes at least 30 minutes. We analyzed a total of 516 patients' datasets.

NIBP (Non-Invasive Blood Pressure) and intravenous anaesthesia data were collected every 3 seconds during surgery, and surgical stages were recorded manually into the EVENT variable. The details of the collected data are presented in Table I. Blood pressure data consisted of NIBP_SBP (Systolic Blood Pressure), NIBP_MBP (Mean Blood Pressure), and NIBP_DBP (Diastolic Blood Pressure). Anaesthesia data consisted of the following parameters for the drugs used during anaesthesia, namely Propofol and Remifentanil: CP (Concentration in Plasma), CE (Concentration in Effect-site), CT (Concentration in Target), RATE (Injection Rate), and VOL (Total Injection Volume).

TABLE I
Data Category

| Category | Variable |
|---|---|
| NIBP | Systolic Blood Pressure, Mean Blood Pressure, Diastolic Blood Pressure |
| PROPOFOL | PROPOFOL_CP, PROPOFOL_CE, PROPOFOL_CT, PROPOFOL_VOL, PROPOFOL_RATE |

| | |
|---|---|
| REMIFENTANIL | REMIFENTANIL_CP, REMIFENTANIL_CE, REMIFENTANIL_CT, REMIFENTANIL_VOL, REMIFENTANIL_RATE |
| EVENT | Recording of events occurring during surgery (e.g. Intubation) |

### *B. Feature Selection*

The surgical stages are identified and labelled in the complete sequence data for each patient. We classified these stages into distinct categories, such as Record start to Induction (0), Induction to Intubation (1), Intubation to Operation start (2), Operation start to Operation end (3), and Operation end to Record end (4), and utilized them as independent features for prediction. Then, using a correlation coefficient, we conducted correlation analysis to assess the linear relationship between two variables. Table II presents the results of the correlation analysis. A coefficient close to 1 indicates a strong positive correlation, while a coefficient close to -1 indicates a strong negative correlation. A coefficient close to 0 suggests no linear correlation between the two variables. In this study, we used the correlation analysis method to identify the variables most strongly associated with predicting anaesthesia complications.

TABLE II
CORRELATION ANALYSIS

| Variable | Correlation Coefficient |
|---|---|
| NIBP_SBP | 1.000 |
| NIBP_MBP | 0.910 |
| NIBP_DBP | 0.689 |
| REMIFENTANIL_CP | 0.212 |
| REMIFENTANIL_CT | 0.192 |
| REMIFENTANIL_CE | 0.187 |
| PROPOFOL_CE | -0.167 |
| PROPOFOL_CT | -0.157 |
| PROPOFOL_CP | -0.103 |
| REMIFENTANIL_RATE | 0.071 |
| PROPOFOL_VOL | -0.014 |
| REMIFENTANIL_VOL | 0.012 |
| PROPOFOL_RATE | 0.007 |

We used a correlation analysis to select features with an absolute correlation coefficient of 0.1 or higher for predicting anaesthesia complications. Based on this method, we selected the following features: surgical stages, NIBP_SBP, NIBP_MBP, NIBP_DBP, REMIFENTANIL_CP, REMIFENTANIL_CT, REMIFENTANIL_CE, PROPOFOL_CE, PROPOFOL_CT, and PROPOFOL_CP. Our anaesthesia complications prediction model will use these features.

### *C. Research Process*

We randomly split the data of 516 patients undergoing general anaesthesia surgery into training and testing sets at an 8:2 ratio. Specifically, we used data from 413 patients for training, while we utilized the data from the remaining 102 patients for testing. Each sample in the training data was observed for 1 minute in the past and used to predict the occurrence of anaesthesia complications 5 minutes after the observation time point using a Transformer model. We constructed the samples by moving one step at a time, as shown in Fig. 1.


Fig. 1 Construction of samples for training data

This study employed a Transformer algorithm to achieve real-time prediction of anaesthesia complications, as depicted in Fig. 2.

Due to handling the lengthy input time of the data used in this study more effectively, we chose a Transformer model with an attention mechanism. This combination has the advantage of being less influenced by input sequence length and processing sequences in parallel while also capturing contextual information to improve prediction accuracy.

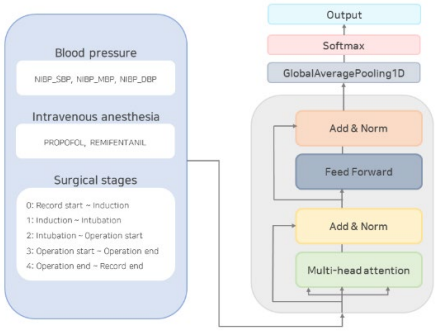



Fig. 2 Research framework

This study includes experiments predicting anaesthesia complications 1 and 5 minutes ahead. In addition, we compare Seq2Seq (Sequence-to-Sequence) and Seq2seq + Attention (Sequence-to-Sequence Model with an Attention Mechanism) to evaluate different models' performance.

## III. RESULTS

In this study, we proposed a Transformer model algorithm to predict blood pressure events during general anaesthesia. We used the F1-Score as the evaluation metric for the classification model to compare the performance. The F1-Score is used when the class distribution is imbalanced, and each class is equally important.

Table III presents the results of predicting blood pressure events 1 minute ahead. The Seq2Seq and Seq2Seq+Attention models achieved an F1-Score of 0.53, while the Transformer model proposed in this study achieved an F1-Score of 0.83.

TABLE III
RESULTS OF PREDICTING 1 MINUTE AHEAD EVENTS

| Metrics / Model | Accuracy | F1-Score | Precision | Recall |
|---|---|---|---|---|
| **Seq2Seq** | 0.85 | 0.53 | 0.53 | 0.53 |
| **Seq2Seq + Attention** | 0.85 | 0.53 | 0.53 | 0.52 |
| **Transformer** | 0.90 | 0.83 | 0.84 | 0.82 |

A confusion matrix is an essential metric for evaluating the performance of a classification model, showing the predicted and actual target variable values in a matrix format. Table IV presents the confusion matrix for our proposed model.

TABLE IV
CONFUSION MATRIX OF PREDICTING 1 MINUTE AHEAD EVENTS USING TRANSFORMER

| Actual \ Predicted | Normal | Hypotension | Hypertension |
|---|---|---|---|
| **Normal** | 230744 | 4312 | 9715 |
| **Hypotension** | 5364 | 16307 | 42 |
| **Hypertension** | 11519 | 46 | 36917 |

Table V evaluates our proposed model's performance for each category. Our proposed model outperformed the previous study [2] in hypotension prediction, achieving an F1 score of 0.77. However, applying the methodology from [2] to our dataset resulted in a lower F1 score of 0.74. Using the previous methodology, it was also challenging to simultaneously predict high and low blood pressure.

TABLE V
RESULTS OF PREDICTING 1 MINUTE AHEAD EVENTS USING TRANSFORMER

| Category \ Metrics | Precision | Recall | F1-Score |
|---|---|---|---|
| **Normal** | 0.93 | 0.94 | 0.94 |
| **Hypotension** | 0.79 | 0.75 | 0.77 |
| **Hypertension** | 0.79 | 0.76 | 0.76 |

When predicting events 5 minutes ahead, our proposed Transformer model outperformed the Seq2Seq and Seq2Seq+Attention models with an overall F1-Score of 0.62, consistent with the results obtained for 1 minute ahead predictions. Table VI presents the performance comparison of each model for the 5 minutes ahead prediction.

TABLE VI
RESULTS OF PREDICTING 5 MINUTES AHEAD EVENTS

| Model \ Metrics | Accuracy | F1-Score | Precision | Recall |
|---|---|---|---|---|
| **Seq2Seq** | 0.83 | 0.43 | 0.49 | 0.42 |
| **Seq2Seq + Attention** | 0.78 | 0.33 | 0.41 | 0.35 |
| **Transformer** | 0.81 | 0.62 | 0.65 | 0.59 |

The confusion matrix for classification 5 minutes ahead using the Transformer model is shown in Table VII, and Table VIII presents the performance evaluation for each category. The F1-Score for Hypotension and Hypertension are 0.42 and 0.55, respectively.

TABLE VII
CONFUSION MATRIX OF PREDICTING 5 MINUTES AHEAD EVENTS USING TRANSFORMER

| Actual \ Predicted | Normal | Hypotension | Hypertension |
|---|---|---|---|
| **Normal** | 216595 | 7102 | 16361 |
| **Hypotension** | 13661 | 7718 | 287 |
| **Hypertension** | 21122 | 247 | 23633 |

TABLE VIII
RESULTS OF PREDICTING 5 MINUTES AHEAD EVENTS USING TRANSFORMER

| Category \ Metrics | Precision | Recall | F1-Score |
|---|---|---|---|
| **Normal** | 0.86 | 0.90 | 0.88 |
| **Hypotension** | 0.51 | 0.36 | 0.42 |
| **Hypertension** | 0.59 | 0.53 | 0.55 |

To enhance the accuracy of our results, we conducted five rounds of validation. We evaluated our findings with a 99% confidence interval, as depicted in Fig. 3. As a result, we can confidently state that our findings are reliable.

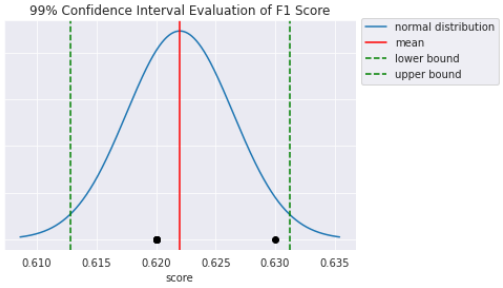



Fig. 3 99% Confidence interval evaluation of F1-Score

## IV. CONCLUSIONS

Developing a reliable and accurate predictive model for blood pressure changes during anaesthesia using non-invasive vital sign data is essential for ensuring patient safety and improving the quality of care. In this study, we propose a Transformer-based model for predicting blood pressure changes during anaesthesia using non-invasive vital sign data has several advantages. First, it only requires additional patient information beyond what is collected during surgery, reducing the potential for errors or inconsistencies arising from such data's absence. Second, the model can make more reliable predictions with fewer data points, which is crucial for preventing complications and saving lives. Finally, it ensures better protection of patient privacy.

Our study has significant implications for improving patient outcomes and advancing anaesthesia management. By enhancing the efficiency and accuracy of real-time prediction of anaesthesia complications, our model can lead to better clinical decision-making and create a safer surgical environment for patients and practitioners. Future research using a Time series GAN (Generative Adversarial Networks) model and incorporating an outlier detection layer will help address the challenge of limited event occurrence data and improve prediction accuracy even further.

## ACKNOWLEDGMENT

This research was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0012724, The Competency Development Program for Industry Specialist) and the Soonchunhyang University Research Fund.

## REFERENCES

[1] Vistisen, S.T., Johnson, A.E.W., and Scheeren, T.W.L., "Predicting vital sign deterioration with artificial intelligence or machine learning," Journal of Clinical Monitoring and Computing, 33(6), pp. 949–951, June 2019.

[2] Jihyun Lee, Jiyoung Woo, Ah Reum Kang, Young-Seob Jeong, Woohyun Jung, Misoon Lee, and Sang Hyun Kim, "Comparative analysis on machine learning and deep learning to predict post-induction hypotension," Sensors 20(16), p. 4575, Aug. 2020.

[3] Anna Tselioudis Garmendia, Ioannis Gkouzionis, Charalampos P. Triantafyllidis, Vasileios Dimakopoulos, Sotirios Liliopoulos, and Marc Chadeau-Hyam, "Towards personalised early prediction of intra-operative hypotension following anaesthesia using deep learning and phenotypic heterogeneity," Pain & Central Nervous System Week, Jan. 2023.

[4] P. -Y. Ko, L. -F. Lai, T. -H. Ku, and Y. -D. Lin, "A CNN approach to predicting non-invasive blood pressure of surgical patients," 2022 IEEE 5th International Conference on Knowledge Innovation and Invention (ICKII), Hualien, Taiwan, 2022, pp. 9–13.

# Sleep GAT : Sleep Score Prediction using Graph Attention Network

YouJung Han[1,*], JiYoung Woo[2]
[1]*Dept. Future Convergence Technology, University of Soonchunhyang, Asan, South Korea*
[2]*Dept. AI and Bigdata, University of Soonchunhyang, Asan, South Korea*
*Contact: jywoo@sch.ac.kr

*Abstract*—**Sleep profoundly influences both physical and mental health, with insufficient sleep not only disrupting daily life but also contributing to various diseases and health issues. Sleep score, a quantitative measure reflecting an individual's sleep state and quality, plays a crucial role in personal health management. This study focuses on predicting sleep scores measured through smartphone and smart watch using graph attention neural network. We adopted the bio signal, activity data and past sleep patterns and built the graph neural network to identify the relationship between those variables. We apply attention mechanism to graphs built from bio signal, activity data and sleep patterns to make the model pay attention to important values. we addressed missing values in daily life data through the use of multivariate time series data and diverse imputation techniques. we also compare our model with machine learning and deep learning algorithm.**

## I. Introduction

Sleep plays a crucial role in human health and daily life, as good sleep significantly contributes to maintaining physical and mental well-being, ultimately improving the quality of everyday life. Predicting sleep scores helps individuals improve their sleep conditions and build healthier lifestyles.

Previous studies have predicted sleep quality based on mobile sensor data and communication data [1], and there have been cases predicting sleep quality using daytime physical activity data [2,5]. Moreover, sleep quality has been predicted using deep learning techniques with wearable sensors [3,4].

This study collected physical activity data, bio signals, and sleep patterns from 41 participants, including Soonchunhyang University's undergraduate and graduate students, using smartphones and smartwatches. unlike previous studies, we conduct analysis through a sleep score that represents one indicator by comprehensively considering various factors such as sleep quality, sleep depth, duration, and number of interruptions. Through this, it helps to find individual sleep patterns and healthy sleep habits, and can help with sleep improvement strategies.

## II. Methodology

### A. Data

We conduct research with 41 participants to collect physical activity data and sleep data using SAMSUNG smartphones and smartwatches from a diverse group of individuals. The smartwatch utilizes its built-in accelerometer, optical heart rate sensor, and electrical heart sensor to track sleep data and track the participants' physical activities during their daily lives using smartphones and smart watch. In Table I, we can find descriptions of the various data types used in this study.

The data types encompass various aspects, such as the user's exercise, floors climbed, heart rate, sleep, sleep stages, daily step count trends, and daily running and walking steps. Within the sleep, variables include mental_recovery,physical_recovery, movement_awakening,sleep_cycle,efficiency,sleep_score and sleep_duration. In total, 25 variables are utilized for the analysis.

The sleep score is formulated and provided by the Samsung Health app in consideration of total sleep time, sleep cycle, movement and wakefulness, physical recovery, and mental recovery factors. Furthermore, Fig. 1 displays the sleep scores for three selected participants, while Table II presents descriptive statistics for the sleep score of 41 participants, including mean, standard deviation, minimum, maximum, and percentile values.

TABLE I
DATA DESCRIPTION

| Category | Data types | Description |
|---|---|---|
| activity | exercise | user's exercise data |
| activity | floors_climbed | user's floors climbed data |
| bio-signal | heart_rate | user's heart rate data |
| sleep pattern | sleep | user's sleep data |
| sleep pattern | sleep_stage | detailed sleep stage for specified sleep data |
| activity | step_daily_trend | user's daily step count trend data |
| activity | pedometer.day_ summary | daily running steps, walking steps for user |

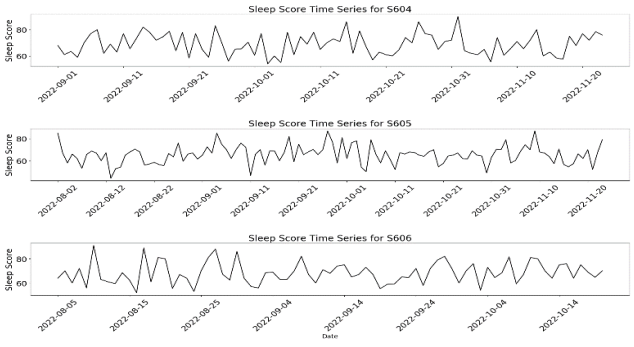



Fig. 1. This figure shows time-series data of sleep scores for three participants.

TABLE III
STATISTIC OF SLEEP SCORE

| Statistic | Value |
|---|---|
| Mean | 65.32 |
| Standard Deviation | 1.67 |
| Minimum | 34.00 |
| 25th Percentile | 59.00 |
| Median | 66.00 |
| 75th Percentile | 72.50 |
| Maximum | 96.00 |

Although no significant patterns or anomalies were found in the time series data of the three participants' sleep scores represented in Fig. 1, further analysis and potential utilization of artificial intelligence techniques could be explored to identify hidden patterns or anomalies in current research.

### *B. Imputation*

Data plays a crucial role in understanding various phenomena, making informed decisions, and improving outcomes in many fields. However, such data often contain missing values, which can lead to biases, reduced statistical power, and compromised validity of analyses. Imputation, a statistical method for handling missing values, has emerged as an important approach to address these issues and enhance the quality and reliability of data analysis.

By performing imputation, researchers can gain more accurate insights, improve the performance of predictive models, and ultimately make better-informed decisions across various domains.

### *C. Time Series Graph*

In this study, we use a method to construct a graph by considering the correlation between features. Correlation coefficients between the variables are calculated, and pairs of interconnected nodes are created as edges when their correlation coefficient exceeds a certain threshold. Fig. 2 serves as an example for the preceding explanation. Highly correlated data is linked within a graph so that a graph-based deep learning model, a GCN (Graph Convolution Network), can capture and learn these relationships.

The correlation coefficient revealed a moderate positive relationship between sleep score and physical recovery (0.52), suggesting higher recovery with better sleep. Sleep cycle positively correlated with physical recovery (0.65), while movement awakening had a strong negative correlation with sleep efficiency (-0.74).

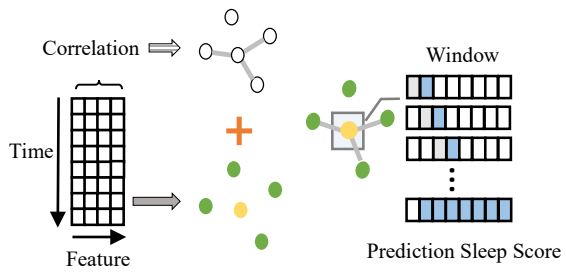



Fig. 2. The graph is constructed by examining the correlation between variables, an edge is created based on the threshold value of the correlation coefficient, and the graph structure is shown when the window is 1 Day

## III. RESULTS

### *A. Imputation*

We conducted experiments to compare the performance of various imputation methods for handling missing values in health data. In these experiments, we introduced 10% missing values to the original dataset by randomly selecting the actual values and replacing them with missing values.

After performing imputation, we compared the imputed values with the actual values to compute the Mean Absolute Percentage Error (MAPE) for each method. The results of the experiments are presented in Table III, which summarizes the MAPE for each method.

TABLE IIIII
RESULTS OF IMPUTATION

| Algorithm | MAPE |
|---|---|
| **Interpolation** | 33.68 |
| **Multiple Imputation** | 30.75 |
| **KNN Imputation** | **24.58** |
| **RNN-based imputation** | 49.85 |

The results reveal that KNN Imputation performed best with the lowest MAPE of 24.58, indicating its accuracy for imputing missing values in our health dataset, while RNN-based Imputation had the highest MAPE of 49.85, suggesting it's not ideal for this dataset. Interpolation and Multiple Imputation showed moderate performance with MAPEs of 33.68 and 30.75. Fig. 3 compares time-dependent interpolation and KNN imputation for handling missing values in time series data. Time-dependent interpolation uses linear estimation, working well for uniformly distributed data. KNN Imputation, on the other hand, fills missing values using nearest-neighbor data, considering data patterns. Fig. 3 shows that KNN Imputation restores missing values more accurately than time-dependent interpolation, even in complex data patterns.


Fig. 3. Comparison of imputed values using Time Interpolation and KNN Imputation: True Data (blue solid line) and Imputed Data (red dashed line)

These findings demonstrate the importance of selecting an appropriate imputation method to achieve better accuracy and performance when dealing with missing data in health datasets.

### *B. Graph Attention Network*

Graph Attention Network (GAT) is an effective deep learning architecture for processing graph-based data, automatically learning the connections and features between nodes while adjusting weights[6]. GAT aggregates information from neighboring nodes to generate new embeddings for each node by utilizing self-attention mechanisms. These mechanisms calculate weights between each node and its neighbors, emphasizing important information and reducing unnecessary information. By implementing GAT, it becomes possible to enhance knowledge extraction and prediction in complex graph structures, offering outstanding performance and efficiency in handling graph data.

### *C. Predict Sleep Score*

In this study, we predict the next day's sleep scores based on the previous 1-day, 3-day, and 7-day values, and we compare the performance of various machine learning algorithms for each prediction horizon. Table IV presents the root mean square error (RMSE) values of each algorithm for the respective prediction horizons.

TABLE IV
RESULTS OF PREDICTING SLEEP SCORE

| Algorithm | RMSE | | |
|---|---|---|---|
| | 1Day | 3Day | 7Day |
| **XGBoost** | 8.12 | 8.25 | 9.37 |
| **Random Forest** | 6.85 | 6.47 | 7.51 |
| **CatBoost** | 6.97 | 6.55 | 7.52 |
| **LSTM** | 7.90 | 7.88 | 9.03 |
| **RNN-GRU** | 7.97 | 7.99 | 7.98 |
| **GCN** | 5.04 | 4.38 | 4.39 |
| **GAT** | **3.83** | 5.02 | 5.61 |

In this study, we aimed to compare the performance of various machine learning and deep learning algorithms for predicting sleep scores using three distinct prediction horizons: 1-day, 3-day, and 7-day. Table IV presents the root mean square error (RMSE) values for each algorithm corresponding to the respective prediction horizons.

Upon closely examining the results, it becomes evident that the Graph Attention Network (GAT) surpasses all other algorithms in performance, demonstrating the lowest RMSE values of 3.83 specifically for the 1-day prediction horizons. This underlines the effectiveness of GAT in accurately predicting sleep scores and highlights its potential for similar tasks. In comparison, other machine learning algorithms such as XGBoost, Random Forest, and CatBoost, as well as deep learning models like LSTM and RNN-GRU, showed higher RMSE values, indicating their relatively lower accuracy in sleep score prediction for the given dataset. Notably, the Graph Convolutional Network (GCN) also demonstrated competitive performance, with RMSE values of 5.04, 4.38, and 4.39 for the three prediction horizons.

By integrating these findings, our research emphasizes the importance of selecting the appropriate algorithm to achieve better performance in predicting sleep scores. The superior performance of GAT demonstrates the potential of graph-based techniques in handling time-series data and suggests that they could be further explored for diverse applications in various data domains.

## IV. CONCLUSION

In this study, we aimed to predict sleep scores using multivariate time series data collected from smartphones and smartwatches and address the problem of missing values in health datasets. Our findings show that KNN imputation is the most accurate way to handle missing values, outperforming other imputation techniques such as interpolation, multiple imputation, and RNN-based imputation.

Among the three windows (1-day, 3-day, and 7-day) considered in this study, it was observed that utilizing data from the previous day yielded the highest performance in predicting sleep scores, with the Graph Attention Network (GAT) model outperforming other models. Furthermore, we explored the potential of combining graphs and multivariate time series data to enhance predictive performance.

Our findings not only contribute to the existing literature on sleep score prediction but also demonstrate the potential of these methodologies for processing everyday life log data and improving the usefulness of smartwatches in personal healthcare. However, the researchers limited the size and diversity of the datasets used in this study, which requires conducting further experiments with more participants and collecting data from different population groups. This is expected to enable the development of a more accurate and stable sleep score prediction model.

## ACKNOWLEDGMENT

This work was supported by the Technology Innovation Program (or Industrial Strategic technology development program, 20015023, Human-following semi-autonomous bed robot for quarantine and transport of infected patients).

## REFERENCES

[1] Xu, Bin, et al. "Will you have a good sleep tonight? sleep quality prediction with mobile phone." 7th International Conference on Body Area Networks. 2012.

[2] Sathyanarayana, Aarti, et al. "Sleep quality prediction from wearable data using deep learning." JMIR mHealth and uHealth 4.4 (2016): e6562.

[3] Arora, Anshika, Pinaki Chakraborty, and M. P. S. Bhatia. "Analysis of data from wearable sensors for sleep quality estimation and prediction using deep learning." Arabian Journal for Science and Engineering 45 (2020): 10793-10812.

[4] Phan, Dinh-Van, Chien-Lung Chan, and Duc-Khanh Nguyen. "Applying deep learning for prediction sleep quality from wearable data." Proceedings of the 4th International Conference on Medical and Health Informatics. 2020.

[5] Kılıç, Ozan, Berrenur Saylam, and Özlem Durmaz İncel. "Sleep Quality Prediction from Wearables using Convolution Neural Networks and Ensemble Learning." *arXiv preprint arXiv:2303.06028* (2023).

[6] Veličković, Petar, et al. "Graph attention networks." arXiv preprint arXiv:1710.10903(2017)

# Development and Evaluation of a Fabric Sensor-Based Balance Assessment Tool

HyoJin Shin[1], JungYeon Kim[1*], and JiYoung Woo[2]
[1] *Dept. ICT convergence, University of SoonChunHyang, Asan, South Korea*
[2] *Dept. AI and Bigdata, University of SoonChunHyang, Asan, South Korea*
*Contact: betterwithme@sch.ac.kr

*Abstract*— **The occurrence of falls can result in severe harm, especially those with unstable balance, such as the patient or the elderly people. This study aimed to develop an effective tool to assess balance that is closely associated with falls.**
**We performed the balance test by tilting a subject in an upright stance using a pressure sensor and built feature set from literature to determine the center of pressure. To evaluate the usefulness of feature set, we built a machine learning model that recognize balance and imbalance utilizing multiple extracted features, including statistical features. The experiment results showed that our features can effectively detect balance and imbalance.**

## I. INTRODUCTION

Falls are a prevalent public health issue and a leading cause of injury, disability, and mortality among the elderly and patients with poor balance. The World Health Organization (WHO) estimates that approximately 646,000 people die each year from falls worldwide, and 37.3 million from non-fatal injuries. Early detection and prevention of imbalances are crucial to avoid such accidents. The balance test using a pressure sensor is a non-invasive and objective method for measuring postural sway, which is an indicator of balance control. The purpose of this study is to detect imbalance by analyzing the pressure sensor data collected during the balance test using machine learning. This study will provide a tool embedding intelligent detection method for residents with risk of falls.

## II. RELATED WORK

Various studies on falls and related problems have been conducted in the past, and studies using the Center of Pressure (COP) have also been conducted. Previous studies have demonstrated the effectiveness of COP parameters in detecting changes in balance [1]. For example, SL Bennett et al. A linear discriminant classifier was used to extract feature points from the COP motion and classify the imbalance with about 90% accuracy [2]. In this study, the accuracy of balance and imbalance detection was improved by integrating not only COP feature points but also statistical characteristics.

Similarly, Marcio R. Oliveira et al. Postural instability was assessed by measuring COP parameters related to agitation and obtained a sensitivity of 78% and specificity of 68% [3]. Ruopeng Sun et al. COP parameters such as COP, mean sway velocity and area, to classify the four fall risk groups with 86% accuracy [4]. Also, Ashirbad Pradhan et al. COP displacement and velocity features were used to classify falls with 84.95% accuracy [5]. In contrast to previous studies, our model expands the capability to identify not only equilibrium and imbalance, but also three distinct states, including the state immediately preceding loss of balance. Moreover, we developed a model that achieves higher accuracy in classifying three specific movements compared to prior research.

## III. METHODOLOGY

### A. Data

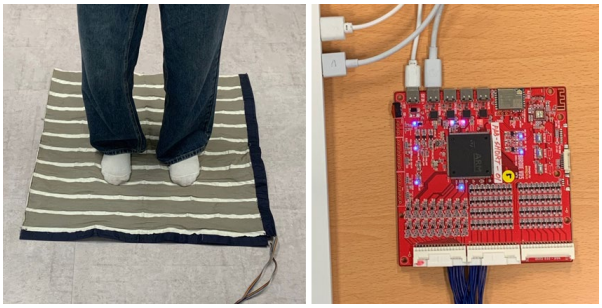


Fig. 1 Experimental Setup

In this study, a data set consisting of three experiments was collected using a cloth press mat in which data were collected at 1 Hz, as shown in Fig 1, to distinguish between balanced and unbalanced.

In the first experiment, 20 healthy participants stood still for 50 seconds with their eyes open. In the second experiment, participants spread their feet 10 cm apart and tilted them 15 degrees up, down, left, and right for 5 seconds [6]. Finally, in the third experiment, 12 healthy subjects were rotated 10 times with the upper body slightly bent, and then closed their eyes and stood still for 50 seconds. Data were collected for each of these experiments.

### B. Preprocess

Due to the small number of participants, a window length of 10 was specified with a 1-second overlap. The second

experiment had less data compared to the first due to the limited time of 20 seconds, while the third experiment had a small number of participants, resulting in a class imbalance issue with less data compared to the first experiment. To alleviate this problem, we applied data augmentation through the SMOTE(Synthetic Minority Over-sampling Technique) algorithm.

To remove sensor data noise, a threshold array was set based on sensor data collected when no pressure was applied, and this was used for primary correction. The pressure sensor values were then calibrated using a upsampling and smoothing with a Gaussian filter [7]. An example of the calibrated data is presented in Figure 2 below.

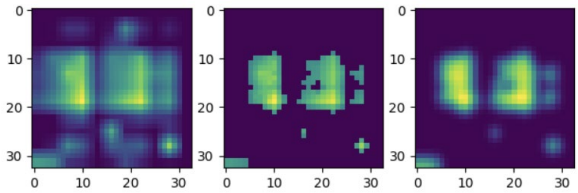


Fig. 2 Data upsampling and smoothing with a Gaussian filter.

The left side of Fig 2 shows the array at the stage following upsampling of the original data but prior to the application of any correction. In the middle panel, the array is depicted after undergoing threshold-based correction. The array on the right-hand side represents the final corrected data that has been subject to the data smoothing process.

### C. Center Of Pressure

The data with noise removed were used to calculate the COP based on the following equation.

$$LatCP(t) = \frac{1}{TS\{v,m\}(t)} \sum y(s) \times CISD\{s\}(t) \quad (1)$$

$$LonCP(t) = \frac{1}{TS\{v,m\}(t)} \sum x(s) \times CISD\{s\}(t) \quad (2)$$

In equations 1 and 2, $t$ represents the time point, $v$ denotes the participant number, and $s$ denotes the sensor number, while $LatCP\{v,\ m\}(t)$ and $LonCP\{v,\ m\}(t)$ respectively represent the x and y coordinates of the COP. The COP coordinates are computed by utilizing the sum of the sensor values ($TS\{v,\ m\}(t)$), the x and y positions of the sensors(x(s),y(s))and along with the individual sensor data (CISD{v,m}{s}(t)) values of the calibration sensors.

### D. Feature Construction

In order to measure the degree of balance based on the calculated COP, various features were extracted [2]. These features include the Sum of Vector Lengths whose movement is above 0.25(SVL), the Number of Acute Angles (NAA), the number of Vector whose movement is above 0.25(NV), the distance between the Start-to-Finish Vector (SFV), the Longest Single COP movement (LSV), as well as statistical features such as variance, standard deviation, and kurtosis of the COP. The calculation formulas for each feature are as follows.

$$l(n) = \sqrt{(x_{n+1} - x_n)^2 + (y_{n+1} - y_n)^2}$$

$$A(l(n)) = \begin{cases} l(n) & if\ l(n) > 0.25\ cm \\ 0 & otherwise \end{cases}$$

$$SVL = \sum_n A(l(n)) \quad (3)$$

In order to implement the SVL feature, we calculated the movement based on the distance l(n) between successive COP coordinates. Here, $x_n$ and $y_n$ represent the x and y coordinates of the COP at each time point, respectively. We considered only those distances with l(n) greater than or equal to 0.25 as meaningful movements for SVL calculation. The SVL value was obtained by summing up all the meaningful distances.

$$theta(n) = \cos^{-1} \frac{l(n) \times l(n+1)}{\|l(n)\| \times \|l(n+1)\|}$$

$$A(theta(n)) = \begin{cases} 1 & if\ theta(n) < 90° \\ 0 & otherwise \end{cases}$$

$$NAA = \sum_n A(theta(n)) \quad (4)$$

In order to implement the NAA (Number of Acute Angles) function, we calculated the angle theta(n) between successive COP coordinates. If theta(n) was an acute angle, it was considered as a sudden movement and the number of such sudden movements was calculated and extracted.

In addition, a feature Percentage of Active Sensors(PAS) was extracted. The calculation formula is as follows [8].

$$PAS(t) = \frac{\sum active\ sensors(t)}{\sum inactive\ sensors(t)} * 100 \quad (5)$$

## IV. RESULTS

### A. Feature Selection

TABLE I
FEATURE IMPORTANCE

| Feature | Score |
|---|---|
| PAS | 437.539 |
| QLR | 260.019 |
| std_x | 165.581 |
| NV | 103.556 |
| std_y | 98.501 |
| var_x | 91.188 |
| SVL | 85.537 |
| SFV | 78.088 |
| LSV | 57.129 |
| var_y | 36.765 |

We preprocessed the data using StandardScaler because the SVM model performs better when the data is in a similar range. Then we used the SelectKBest class from the scikit-learn library to select the most important features of the model. We specified K=10, and the top 10 features were selected, which include PAS, QLR, STD_X, NV, STD_Y, VAR_X, SVL, SFV, LSV, and VAR_Y. The importance of these features is consistent with that shown in TABLE1.

### *B. Data Augmentation*

To increase the relatively small amount of data due to the short experiment time, the SMOTE algorithm was used for data augmentation. The number of Original dataset and After augmentation dataset is shown in Table II below.

TABLE Ⅱ
THE NUMBER OF DATA BEFORE AND AFTER DATA AUGMENTATION

| | Test1 | Test2 | Test3 |
|---|---|---|---|
| Original dataset | 77 | 31 | 46 |
| After augmentation | 77 | 77 | 77 |

### *C. Classification*

The SVM model was used to classify each test. SVM is a model that defines the decision boundary, which is the dividing line for classification. Therefore, when a new point that has not been classified appears, the model performs the classification task by checking which side of the boundary it belongs to.

After performing data scaling, feature selection, and data augmentation, we built an SVM model and obtained the classification results for each experiment, which are presented in Table III below.

TABLE Ⅲ
ACCURACY AND F1-SCORE

| | Accuracy | F1-score |
|---|---|---|
| Test1 vs Test2 vs Test3 | 0.974 | 0.968 |
| Test1 vs Test2 | 0.926 | 0.925 |
| Test1 vs Test3 | 0.968 | 0.964 |
| Test2 vs Test3 | 1.000 | 1.000 |

The extracted features, utilizing the ratio of the center of pressure and the active sensor, demonstrated a high level of accuracy in distinguishing the three experiments, achieving a rate of 97.4%. Additionally, the proposed features distinguished each individual experiment with a high degree of accuracy: 92.6% accuracy for tests 1 and 2, 96.8% accuracy for tests 1 and 3, and 100% accuracy for tests 2 and 3. These results suggest that the proposed features are effective tools for accurately distinguishing between individuals who have a stable balance, those who are on the verge of losing balance, and those who are in the process of falling.

## V. CONCLUSIONS

In this study confirmed that the extracted features are an effective tool for distinguishing between balanced and unbalanced states. The study achieved higher accuracy in distinguishing various states of balance compared to previous studies. However, the limited sample size raises concerns about the generalizability of the results. Furthermore, simulations of impending loss of balance based solely on data from healthy adults are not representative of all fall patients. Therefore, further experiments involving a more diverse range of fall patients are crucial for a comprehensive understanding of the epidemiology and underlying causes of falls. Future research aims to enhance the accuracy and efficiency of balance classification for a wider range of patients by utilizing data obtained from additional experiments.

## ACKNOWLEDGMENT

This work was supported by the Technology Innovation Program (or Industrial Strategic technology development program, 20015023, Human-following semi-autonomous bed robot for quarantine and transport of infected patients).

## REFERENCES

[1] Flavien Quijoux, Aliénor Vienne-Jumeau, François Bertin-Hugault, Philippe Zawieja , Marie Lefèvre , Pierre-Paul Vidal , Damien Ricard : Center of pressure displacement characteristics differentiate fall risk in older people: A systematic review with meta-analysis. ScienceDirect(2020), 64

[2] Bennett, S. L., Goubran, R., Rockwood, K., & Knoefel, F. : Distinguishing between stable and unstable sit-to-stand transfers using pressure sensors. IEEE International Symposium on Medical Measurements and Applications Proceedings (2014), 1-6.

[3] Oliveira, M. R., Vieira, E. R., Gil, A. W. O., Fernandes, K. B. P., & Silva : One-legged stance sway of older adults with and without falls, PloS one(2018), 13(8)

[4] Sun, R., Hsieh, K. L., & Sosnoff, J. J : Fall risk prediction in multiple sclerosis using postural sway measures: A machine learning approach. Scientific Reports.(2019), 9, 16154.

[5] Pradhan, A., Oladi, S., Kuruganti, U., & Chester, V : Classification of Elderly Fallers and Non-fallers Using Force Plate Parameters from Gait and Balance Tasks. SpringerLink(2020).

[6] Brauer, S. G., Burns, Y. R., & Galley, P. : A prospective study of laboratory and clinical measures of postural stability to predict community-dwelling fallers. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences(2000), 55(8), M469-M476.

[7] Cheng, J., Sundholm, M., Zhou, B., Hirsch, M., Mikhelson, I. V., & Virkki, J. : Smart-surface: Large scale textile pressure sensors arrays for activity recognition. Pervasive and Mobile Computing (2016), 28, 42-56.

[8] Bennett, S., Ren, Z., Goubran, R., & Knoefel, F. : In-bed mobility monitoring using pressure sensors. IEEE Transactions on Biomedical Engineering (2015), 62(3), 933-940

# ESTCNN-MultiHead Attention Model for Driver Drowsiness Detection Using EEG Data

Minseop Lee[1], JiYoung Woo[1*]
[1]*Dept. Future Convergence Technology, University of Soonchunhyang, Asan, South Korea*
*Contact: jywoo@sch.ac.kr

*Abstract*— **Driver drowsiness is one of the major causes of traffic accidents, and accurately identifying driver drowsiness is important. In this study, we propose a method to detect driver fatigue based on single-channel EEG data. To achieve this, we apply the continuous wavelet transform to the EEG data to extract features in the time-frequency domain. Subsequently, we develop a hybrid model composed of a multi-head attention mechanism and an EEG-based Spatio-Temporal Convolutional Neural Network (ESTCNN). The ESTCNN is an advanced architecture specifically designed to capture temporal characteristics from frequency data more effectively than traditional CNNs. By combining the multi-head attention mechanism and the ESTCNN, the hybrid model can detect subtle spatial-temporal changes in the EEG signal more effectively. As a result, we obtained a more stable model with higher accuracy compared to a standalone ESTCNN model. This study offers an accurate and efficient means of detecting fatigue in the field of EEG-based driving fatigue classification.**

## I. INTRODUCTION

Drowsiness is a common issue faced by most people and can result in negative consequences such as decreased productivity, impaired cognitive function, and increased risk of accidents. According to the National Highway Traffic Safety Administration (NHTSA), over 70,000 police-reported accidents were caused by drowsy driving between 2009 and 2013 [1]. Early detection of drowsy driving is crucial in reducing traffic accidents and creating a safe driving environment. Using EEG signals for drowsiness detection is a promising approach that is non-invasive and allows for monitoring of brain activity. Traditionally, multi-channel EEG systems have been used for drowsiness detection in various situations, but they were inconvenient to use and could be costly. Recently, researchers have been exploring the potential of single-channel EEG data in various applications, including drowsy driving detection. However, developing an efficient and accurate drowsy driving detecting model using single-channel EEG data remains a challenge. Lal et al. found a correlation between theta and delta wave activity and drowsiness in their study [2]. Chin-Teng Lin et al. used a 30-channel EEG dataset and performed independent component analysis, discovering that the C3-C4 and P3-P4 channels, located centrally in the scalp topography, were closely related to drowsiness [3]. Balam et al. compared a deep learning (DL) approach using multichannel EEG data collected from the Cz-Oz channel with other feature-based approaches, concluding that the DL approach yielded results at least 3% better [4]. Yingying et al. augmented single-channel EEG data using GANs, applied continuous wavelet transform (CWT) to generate power spectral density (PSD) features, and achieved 98% accuracy on their dataset using an LSTM model [5]. Chen et al. used network analysis to examine the differences between alert and drowsy states for each brainwave frequency band (delta, theta, alpha, beta), observing a clear difference between delta and theta waves in alert and drowsy states [6]. Z. Gao et al. proposed an EEG-based Spatio-Temporal CNN (ESTCNN) model structure, which was more effective in processing data containing temporal information [7]. J. Cui et al. transformed single-channel EEG data using Welch's method into a vector containing relative information for each brainwave frequency band and trained a CNN model [8]. In this study, we focused on developing a robust drowsiness detection model using single-channel EEG data. We used the structure of the ESTCNN model. We compared the performance of the original ESTCNN model trained on raw EEG data, the same model using Morlet CWT applied to the EEG data, and the ESTCNN model combined with the multi-head attention. Through this approach, we proposed a more effective drowsiness detection model by incorporating multi-head attention, which enables the comprehensive learning of the relationships in both the frequency domain and time domain for each input, leveraging the application of CWT on EEG data. By incorporating this approach, we were able to consider individual differences and enhance the overall performance of the system. This proposed model offers a more effective drowsiness detection capability.

## II. METHODS

In this section, we describe the key methodologies used in the research, including the dataset employed, the data pre-processing, and the model structure.

### *A. Dataset Employed*

In this study, we used the dataset collected by Z. Cao et al., who employed a virtual-reality dynamic driving simulator with a 32-channel EEG recording system to measure brain dynamics and behavioral performance in a sustained-attention driving task [9]. EEG datasets were collected from 27 voluntary participants with a total of 62 datasets. A virtual-reality dynamic driving simulator was used to perform a sustained-attention driving task involving lane departure events, and the participants' reaction time was measured. This data does not have ground truth information on drowsiness. Therefore, we labelled the data based on the participant's reaction time distribution using the method proposed by J. Cui et al., who also used the same data for their research [8].

### B. Data Pre-processing

The raw data was meticulously pre-processed by Z. Cao et al., with band-pass filters applied, and effectively removing artifacts caused by eye blinks and eye movements [10]. We extracted data only from the Oz channel, since N.R. Pal et al. reported that the Oz channel showed a more distinct difference between drowsiness and alertness compared to other channels [11]. We then applied the Morlet CWT with scales of 0.5, 1, 2, 4, 8, 16, and 32 to include information from each brainwave frequency band.

### C. Model structure

In this study, three models are used: the Compact CNN as a baseline model, which is the CNN model employed by J. Cui et al. [8], the ESTCNN model, and the ESTCNN model combined with multi-head attention. The Compact CNN is a model proposed in a prior study conducted with the same data and is utilized as a baseline model in this research [8]. The ESTCNN model is applied since the data extracted from the Oz channel is one-dimensional time-domain data, and the ESTCNN can manage temporal information more efficiently. Additionally, with the application of the CWT, the data is transformed into a two-dimensional representation in both frequency and time domains. Due to these data characteristics, the ESTCNN model combined with multi-head attention is employed. This model is expected to manage the interactions and complex patterns between the frequency and time domains more efficiently. The Compact CNN takes a one-dimensional signal as input and applies 32 1D convolutional filters to output a multivariate signal of 32 dimensions. Then, a batch normalization layer and an ELU activation function are applied, followed by a global average pooling layer and a softmax function, as shown in Fig. 1.

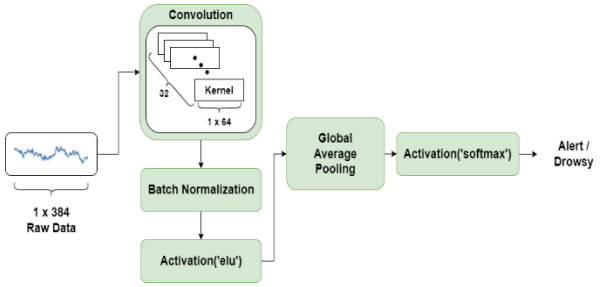



Fig. 1 The structure of the compact CNN model

The ESTCNN consists of three stages: the temporal dependencies extraction stage with three core blocks, the spatial features fusion stage, and the classification stage. Each core block consists of three convolutional blocks and a pooling layer, with max pooling applied in the first and second core blocks and average pooling in the last core block. The core blocks extract features from various time intervals in the input signal and combine them into different channels to form high-level features. After passing through the three core blocks, the features are flattened and fused spatially before being fed into a fully connected layer for classification. Our proposed ESTCNN model combined with multi-head attention uses multi-head attention after the temporal dependencies extraction stage to learn the relationship between the time domain and the frequency domain. Multi-head attention, which is a prominent component of the 'Transformer' architecture primarily used in the field of natural language processing, allows the model to simultaneously grasp the interdependencies among words in a sentence from various perspectives. This mechanism enables the model to learn multiple relationships and discern the importance of words in context. By applying it to EEG data with Continuous Wavelet Transform (CWT), the model can learn the relationships of each input in a complex manner, encompassing both the frequency domain and the time domain. We recommend applying it to transformed EEG data such as FFT (Fast Fourier Transform) and STFT (Short-Time Fourier Transform) rather than raw EEG data without any transformations. FFT is a widely used technique for converting a time-domain signal into its frequency-domain representation, while STFT represents a compromise between time and frequency resolutions by analyzing a signal over short periods of time. CWT (Continuous Wavelet Transform) provides a high-resolution representation of both time and frequency by analyzing a signal at different scales. The reason for this recommendation is that by applying multi-head attention, we expect to manage the interactions and complex patterns between the frequency and time domains more efficiently. Fig. 2 and Fig. 3 illustrate the structures of ESTCNN and our proposed model, respectively.

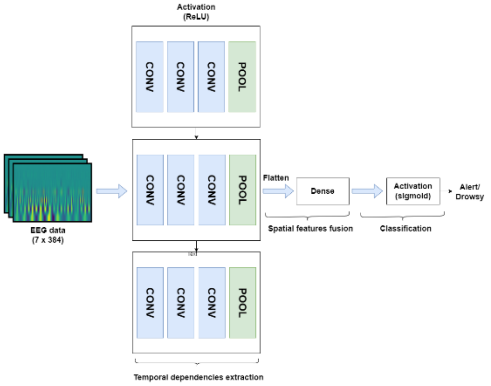



Fig. 2 The structure of the ESTCNN model

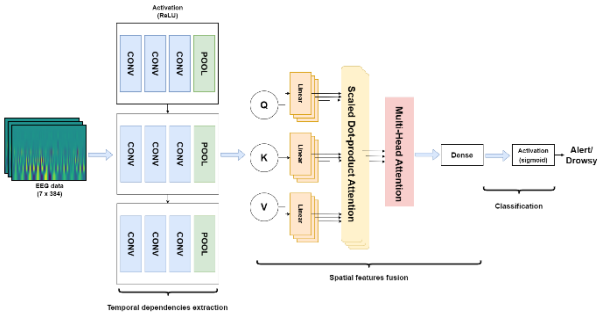



Fig. 3 The structure of proposed model

## III. Results

Using data collected from 27 subjects, we conducted Leave-One-Out Cross Validation based on subject ID using only data from subjects 1 to 11 to measure the accuracy of Compact CNN, ESTCNN, and the proposed model. For ESTCNN, we conducted separate evaluations for 1D EEG data without applying CWT and 2D EEG data with CWT applied. The results are shown in Table 1.

TABLE I
Comparison Of Accuracy Among Each Model

| Subject | Compact CNN | ESTCNN (1D) | ESTCNN (2D) | Proposed model |
|---|---|---|---|---|
| 1 | 77.45 | 82.87 | 85.10 | **86.17** |
| 2 | 52.80 | 57.72 | 57.87 | **59.09** |

| | | | | |
|---|---|---|---|---|
| 3 | 63.47 | 62.79 | **65.06** | 64.66 |
| 4 | 76.22 | 66.21 | 63.51 | **78.37** |
| 5 | 76.52 | 78.48 | 74.19 | **83.03** |
| 6 | 77.11 | 83.37 | 84.21 | **86.74** |
| 7 | 67.35 | **69.41** | 68.82 | 64.11 |
| 8 | **71.93** | 70.22 | 68.56 | 68.40 |
| 9 | 88.25 | **89.74** | 85.54 | 88.66 |
| 10 | 81.67 | 87.39 | 87.59 | **87.77** |
| 11 | 72.65 | 75.39 | **76.90** | 76.28 |
| Mean | 73.22 | 74.89 | 74.30 | **76.66** |

Among the four cases compared, the proposed model with multi-head attention showed the best performance with an average accuracy of 76.66%. Considering that the proposed model using multi-head attention improved performance compared to other models that did not use multi-head attention, using multi-head attention helped to extract better features by considering the relationships between input tensors.

## IV. Conclusions

In conclusion, the proposed model with multi-head attention outperformed the other models in terms of accuracy. While applying CWT and transforming the data into 2D format may result in some loss of temporal information, the proposed model addressed this issue by utilizing multi-head attention to analyze the data while considering the inter-relationships between input tensors for better feature extraction. Although this study was limited to the EEG domain, further research is needed to investigate whether this approach can be effective in other domains. In addition, the data used in this study lack ground truth labels for drowsy and alter states. Therefore, labelling for drowsy and alter was done by J. Cui et al., based on statistical techniques using reaction time (RT) for randomly generated lane departure events [8]. This introduces a limitation to the study that labelling has not been made according to actual sleepiness.

## Acknowledgment

This research was supported by “Regional Innovation Strategy (RIS)” through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE)(2021RIS-004).

## References

[1] N.H.T.S. Administration, “2018 Fatal Motor Vehicle Crashes: Overview,” 2019.

[2] S. K. Lal and A. Craig, “A critical review of the psychophysiology of driver fatigue,” in Biological Psychology, vol. 55, no. 3, 2001, pp. 173-194.

[3] Chin-Teng Lin, Ruei-Cheng Wu, Sheng-Fu Liang, Wen-Hung Chao, Yu-Jie Chen and Tzyy-Ping Jung, "EEG-based drowsiness estimation for safety driving using independent component analysis," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 52, no. 12, pp. 2726-2738, Dec. 2005.

[4] V. P. Balam, V. U. Sameer and S. Chinara, "Automated classification system for drowsiness detection using convolutional neural network and electroencephalogram," in IET Intelligent Transport Systems, vol. 15, no. 4, pp. 514-524, 2021.

[5] Y. Jiao, X. Liu, Z. Ma, Y. Zhao and C. Chen, "Driver sleepiness detection from EEG and EOG signals using GAN and LSTM networks," Neurocomputing, vol. 408, pp. 100-111, Jan. 2020,

[6] J. Chen, Y. Liu, J. Ma and G. Wang, "Graph analysis of functional brain network topology using minimum spanning tree in driver drowsiness," Cognitive Neurodynamics, vol. 12, pp. 569-581, 2018.

[7] Z. Gao et al., "EEG-Based Spatio–Temporal Convolutional Neural Network for Driver Fatigue Evaluation," in IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 9, pp. 2755-2763, Sept. 2019, doi: 10.1109/TNNLS.2018.2886414.

[8] J. Cui, Z. Lan, Y. Liu, R. Li, F. Li, O. Sourina and W. Müller-Wittig, "A compact and interpretable convolutional neural network for cross-subject driver drowsiness detection from single-channel EEG," in Methods, vol. 202, pp. 173-184, 2022.

[9] Z. Cao, M. Chuang, J. T. Kingand C.-T. Lin, “Multi-channel EEG recordings during a sustained-attention driving task (raw dataset)”. figshare, 05-Feb-2019, doi: 10.6084/m9.figshare.6427334.v5.

[10] Z. Cao, C.-T. Lin, C.-H. Chuang, K.-L. Lai, A. C. Yang, J.-L. Fuh, and S.-J. Wang, "Multi-channel EEG recordings during a sustained-attention driving task," Scientific Data, vol. 6, no. 1, p. 19, Jan. 2019. DOI: 10.1038/s41597-019-0023-9.

[11] N. R. Pal, Z. Sankari, Y. Kabir, and A. Bhattacharyya, "EEG-based subject-and session-independent drowsiness detection: An unsupervised approach," EURASIP J. Adv. Signal Process., vol. 2008, no. 1, art. no. 519480, 2008.

# An Object Defect Detection System using YOLOv7

Heebum Kim[1,*], Byeong-Gwon Kang[1], and Yunyoung Nam[1]
[1]*Dept. of ICT Convergence, Soonchunhyang University, Asan, South Korea*
*Contact: kbum0107@naver.com, phone +82-10-9925-6003

*Abstract*— **Defect and anomaly detection are crucial for ensuring product quality and safety in various industrial fields. However, traditional detection methods such as manual inspection and rule-based systems have limitations in terms of efficiency and accuracy. To overcome these challenges, the YOLO algorithm is utilized to detect defects and anomalies in real-time. The proposed system in this paper is designed to classify the defects and anomalies using the YOLOv5 and YOLOv7 model. To improve the accuracy and robustness of the system, data augmentation techniques such as rotation, scaling, and flipping are employed during the training process. The experimental results show that the proposed system achieves high accuracy and real-time performance in detecting various types of defects and anomalies compared to conventioned methods.**

## I. INTRODUCTION

There has been increasing attention to defect detection in recent decades. Previously, the YOLOv1 model was used to perform object detection and minimize errors, but still required a significant amount of computation to extract relevant information from objects. This paper aims to apply the YOLOv5 and YOLOv7 models to achieve better performance in object detection, and to compare and analyze which model is more suitable. It is essential to develop effective algorithms to detect and recognize defects with enough accuracy. This paper aims to explore and compare different approaches to achieve this goal[1].

Object detection is a critical task in computer vision that involves identifying and locating objects within an image or video. The primary stages of object detection comprise feature extraction [2], feature processing [3-5], and object classification [6]. Traditional methods for object detection have achieved excellent performance, and can be categorized into four aspects: bottom feature extraction, feature coding, feature aggregation, and classification. Feature extraction is a vital step in the object detection and recognition process [7], where more redundant information can be modeled to enhance performance beyond previous point-of-interest detection techniques. Examples of these techniques include the scale-invariant feature transformations (SIFT) [8].

Object detection and anomaly detection are critical topics in the field of computer vision and are widely used in various industries. However, existing object detection and anomaly detection systems still have limitations. For example, the detection rate can vary significantly depending on factors such as the size, position, and brightness of the object.

Therefore, the proposed system aims to accurately detect objects, unlike traditional systems, and provides a differentiated approach that can perform both object and anomaly detection. The main goal of this paper is to propose the system, evaluate its performance, and demonstrate its superiority in terms of accuracy and speed compared to existing systems. To achieve this, the performance of the proposed system is evaluated on various datasets of objects and defects, and compared with state-of-the-art object detection and anomaly detection systems. Finally, the paper discusses the practical applicability of the proposed system in various industries and highlights its contribution to the development of object detection and anomaly detection technologies.

## II. METHODS

To detect defects in objects, various types of data were prepared and labelled by marking them with rectangular areas. The object targets were trained for different types of defects such as cracks, scratches, wrinkles, and dents. In the first step, hundreds of images were labelled by loading data files gathered from websites such as Roboflow. In the second step, model performance was tested using function loss rate, precision, recall, and F1-Score. Finally, to verify actual prediction accuracy, images saved in the deep learning folder were loaded and recognition rate was checked.

### *A. Dataset preparation*

The dataset is crucial for training and validating machine learning algorithms. In this paper, we constructed a dataset of products produced in a specific industry, as shown in Fig. 1. The dataset contains various types of defects, as well as normal products. During the construction of this dataset, we collected data considering various environments that occur on actual production lines.

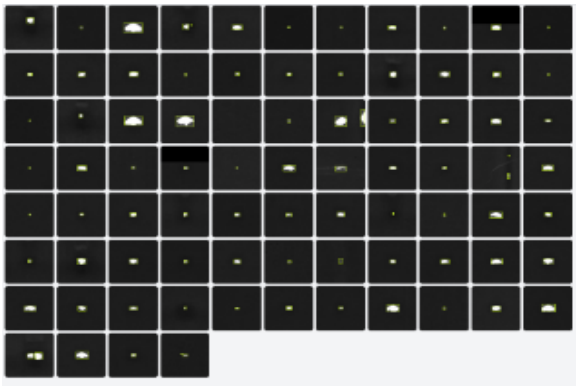


Fig. 1 The process of labelling for defect recognition

*B.* Model training and evaluation

Model training and evaluation are the most important parts of the methodology proposed in this paper. We used the YOLOv5 algorithm or YOLOv7 algorithm to train the model, as shown in Fig. 2, and evaluated the performance of the model using evaluation metrics such as intersection over precision, recall, and F1-Score. Moreover, we optimized the hyperparameters of the model through various validation methods.

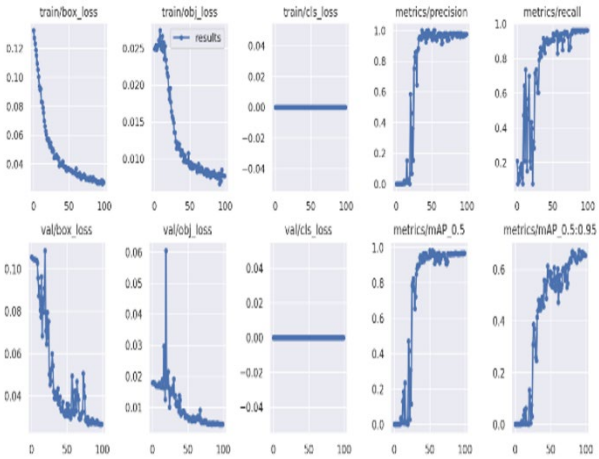



Fig. 2 Train and validation loss, precision, and recall

*C.* System implementation

The system proposed in this paper is implemented to be used in actual industrial settings. To achieve this, we implemented the proposed algorithm as a software system that can be used in actual environments. The system as shown in Fig. 3, inspects products in real-time and is designed to identify and classify detected defects. Additionally, the system uses parallel processing technology for large-scale data processing.

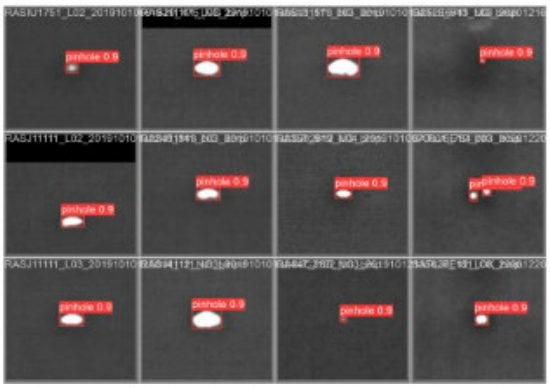


Fig. 3 Actual defect detection rate of the prediction model

## III. RESULTS

After labelling defects such as boundline, crack, crater, craterpinhole, dent, island, line, pinhole, scratch, tape, wrinkle, and background, the ratio of TP, FP, FN, and TN was calculated. The YOLO algorithm correctly identified over 90% of the crack, crater, dent, pinhole, and tape defect images as positive. However, some images were misidentified as a different defect or as background. For example, 58% of boundline defect images were misidentified as background, and 60% of wrinkle defect images were also misidentified as background. In addition, although the crack defect was identified accurately at a rate of 90%, 10% of the images were not recognized, and for line defect images, while 80% were correctly identified as line, the remaining 20 % were identified as background. These results indicate that additional training is necessary to improve model performance, and to achieve this, tasks such as collecting more data, applying various augmentation techniques, rotating images, and resizing them are required.

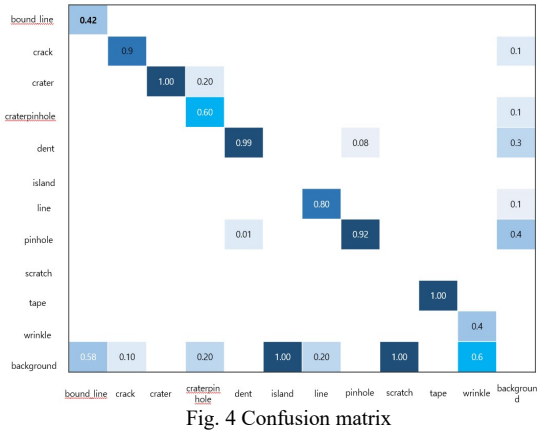



Fig. 4 Confusion matrix

In Figure 5, the validation recognition rate was between 0.5 and 0.7, indicating that the model was overfitting and not generalizing well on the validation data. To address this issue, one solution is to augment the dataset by transforming the data through techniques such as image rotation, translation, resizing, and noise adjustment.

Additionally, improving the accuracy of labelling through review and modification can also help improve the performance of the model. To improve the model's ability to generalize and learn specific types of defects accurately, it is essential to collect more data that represents a wider range of cases.

Fig. 5 Recognition rate about crater

## IV. CONCLUSION

In this paper, we propose a system utilizing YOLO object detection technology for detecting defects and anomalies in the industrial field, and confirm through experimental results that this model shows higher accuracy and performance than other detection models. These results demonstrate the suitability of the proposed system for practical application in the industrial field, and will help in building an accurate and efficient system for quickly detecting product defects and anomalies in the manufacturing process.

Furthermore, this system has room for further development, and accuracy and efficiency can be further improved by utilizing better AI technology and larger datasets. Such advances will be of great help in building more advanced defect and anomaly detection systems in the industrial field. Therefore, the results of this paper are expected to promote research and development in the industrial field, and contribute to the construction of high-performance defect and anomaly detection systems in the industrial field. Therefore, the results of this paper are expected to promote research and development in the industrial field, and contribute to the construction of high-performance defect and anomaly detection systems that can be applied in actual industrial sites.

## ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2022H1D8A3038040)

## REFERENCE

[1] Ahmad, Tanvir, et al. "Object detection through modified YOLO neural network." *Scientific Programming* 2020 (2020): 1-10.

[2] Tiwari, Aastha, Anil Kumar, and Goswami Mansi Saraswat. "Feature extraction for object recognition and image classification." *International Journal of Engineering Research & Technology (IJERT)* 2.10 (2013): 2278-0181.

[3] Yan, Junjie, et al. "The fastest deformable part model for object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014.

[4] Dean, Thomas, et al. "Fast, accurate detection of 100,000 object classes on a single machine." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2013.

[5] Viola, Paul, and Michael J. Jones. "Robust real-time face detection." *International journal of computer vision* 57 (2004): 137-154.

[6] Sun, Da-Wen, ed. *Computer vision technology for food quality evaluation*. Academic Press, 2016.

[7] Wang, Eric Ke, et al. "Deep fusion feature based object detection method for high resolution optical remote sensing images." *Applied Sciences* 9.6 (2019): 1130.

[8] Lowe, David G. "Distinctive image features from scale-invariant keypoints." *International journal of computer vision* 60 (2004): 91-110.

# Enhancing Solar Irradiance Forecasting with Location-Dependent Activation Functions in a Hybrid Deep Learning Approach

Dayeong So[1], Jihoon Moon[1,2,*], and Seungmin Rho[3]

[1] *Department of ICT Convergence, Soonchunhyang University, Asan, South Korea*
[2] *Department of AI and Big Data, Soonchunhyang University, Asan, South Korea*
[3] *Department of Industrial Security, Chung-Ang University, Seoul, South Korea*
*Contact: jmoon22@sch.ac.kr, phone +82-41-530-4956

*Abstract*—**A precise solar irradiance prediction model is essential due to the region's increasing demand for renewable energy. Despite the hybrid deep learning model's exceptional performance in accurate irradiance forecasting, constructing it demands considerable expertise due to the many hyperparameters involved. The study highlights the critical role of activation functions in deep learning models, as they determine each neuron's output and significantly affect the model's performance. To do this, we first developed an innovative hybrid deep learning model for precise short-term solar irradiance predictions in Jeju Island, South Korea. The model processed data from the Korea Meteorological Administration for two Jeju Island regions by combining gated recurrent unit, temporal convolutional networks, and self-attention mechanisms. The study examined the effectiveness of various location-dependent activation functions, such as rectified linear unit (ReLU), scaled exponential linear unit (SELU), and Leaky ReLU, aiming to identify the most suitable activation function for the hybrid deep learning model. The results revealed that the optimal activation function was location-dependent, with SELU consistently outperforming the other functions in both regions. This study enhanced the hybrid deep learning model's applicability for real-world energy generation planning and smart energy management systems by understanding the importance of location-specific activation functions.**

## I. INTRODUCTION

The depletion of fossil fuels and increasing concerns about environmental pollution have intensified the global pursuit of low-carbon, sustainable renewable energy generation [1]. Solar power has emerged as an eco-friendly solution for converting solar energy into electricity, constituting a significant proportion of the overall renewable energy supply. Declining photovoltaic (PV) installation costs, advancements in renewable energy technology, and enhancements in solar module performance have collectively contributed to increased energy generation per unit area, making solar power an efficient energy supply solution for countries with constrained energy generation capacity. In this context, accurate solar irradiance prediction models have become essential due to the growing demand for renewable energy.

Recent studies have focused on improving solar irradiance prediction models [2]–[5]. Ma et al. [2] proposed a self-adaptive model that employed the kernel density estimation algorithm to calculate PV power prediction intervals. They introduced a dynamic interval prediction method with a seasonal distribution feature of solar power prediction errors, which improved the prediction interval coverage probability by 5%. Bhatt et al. [3] developed a deep hybrid model that combined convolutional neural network (CNN) and long short-term memory (LSTM) techniques. This hybrid model outperformed standalone models in predicting ultra-short-term solar irradiance up to six steps ahead at 15-minute intervals, demonstrating its effectiveness for real-time microgrid energy management systems (EMSs). Chang et al. [4] introduced the traditional encoder single deep learning (DL) model for short-term prediction of solar power systems and achieved a 27% increase in accuracy compared to a previously used prediction model. Finally, Luo and Zhang [5] proposed a two-stage learning strategy and an integrated adaptive LSTM model for predicting one-day-ahead PV power generation, outperforming the offline-LSTM model in all cases by detecting concept drift in new data.

Despite the significant advancements in solar irradiance prediction models through these previous studies, certain limitations persist. Most notably, the existing research needs to sufficiently address the role of location-specific activation functions within DL models, which can significantly impact their performance [6]. Moreover, many previous studies relied on singular DL techniques, potentially overlooking the benefits of hybrid models. This study aims to develop an innovative hybrid DL model for precise short-term solar irradiance predictions in Jeju Island, South Korea. Our study underscores the critical role of activation functions in DL models, as they determine each neuron's output and considerably impact the model's performance.

The model integrates gated recurrent units (GRUs), temporal convolutional networks (TCNs), and self-attention mechanisms to address this challenge. It processes data from the Korea Meteorological Administration (KMA) for two Jeju Island regions. It examines the effectiveness of various location-dependent activation functions, including rectified linear unit (ReLU), scaled exponential linear unit (SELU), and Leaky ReLU, to identify the most suitable activation function for the hybrid DL model.

The contributions of our study are as follows:

- We enhanced the understanding of the critical role of location-specific activation functions in DL models for solar irradiance prediction.

- We demonstrated the feasibility and effectiveness of using a hybrid DL approach to address the challenges of renewable energy generation prediction.
- We highlighted the potential of integrating advanced DL techniques with decentralized EMSs for real-world energy generation planning and intelligent EMSs.

The remainder of the paper is organized as follows: Section II outlines the data preprocessing, including data acquisition and feature engineering; Section III details the forecasting model construction, including the hybrid DL model and the various activation functions; Section IV discusses the results and analysis of the model's performance with different activation functions; and finally, Section V concludes the paper with a summary of the findings and suggestions for future research directions.

## II. Date Preprocessing

In this study, we aimed to construct a solar irradiation forecasting model using date/time, meteorological data, and historical solar irradiation data provided by the Korea Meteorological Administration (KMA). We focused on two regions, Ildo-1 dong and Gosan-ri, located on Jeju Island, the largest island in South Korea, which is actively implementing various measures to transition into a smart island by shifting from conventional fossil fuels to renewable energy sources (RESs) [7]. The data collection period spanned eight years, from 2011 to 2018, and occurred between 8 a.m. and 6 p.m. During this time, we collected data on temperature, humidity, wind speed, solar irradiation, and other meteorological observation data such as soil temperature, total cloud volume, ground-surface temperature, and sunshine amount. However, we limited our analysis to sky condition, temperature, humidity, and wind speed, provided by KMA's short-term weather forecasts, as shown in Figure 1. Considering these two regions' solar irradiance and weather conditions, we aimed to assess the applicability of PV systems in South Korea, given Jeju Island's commitment to energy independence through renewable energy resources.

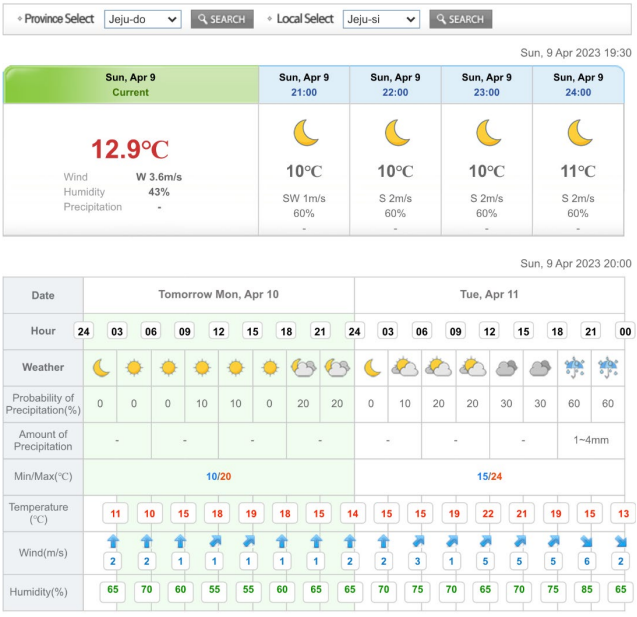



**Figure 1.** KMA's short-term weather forecasts.

TABLE 1.
COMPREHENSIVE LIST OF INPUT VARIABLES (IVS) AND FEATURES.

| IV # | Description (Data Type) | IV # | Description (Data Type) |
|---|---|---|---|
| $Date_x$ | $Date_x$ (continuous) | W3 | Mostly cloudy (binary) |
| $Date_y$ | $Date_y$ (continuous) | W4 | Cloudy (binary) |
| T8 | 8 a.m. (binary) | Temp | Temperature (continuous) |
| T9 | 9 a.m. (binary) | Humi | Humidity (continuous) |
| T10 | 10 a.m. (binary) | WS | Wind speed (continuous) |
| T11 | 11 a.m. (binary) | $D1_{Day\_sin}$ | $Date_{x,D-1}$ (continuous) |
| T12 | 12 p.m. (binary) | $D1_{Day\_cos}$ | $Date_{y,D-1}$ (continuous) |
| T13 | 1 p.m. (binary) | $D1_{W1}$ | $Clear_{D-1}$ (binary) |
| T14 | 2 p.m. (binary) | $D1_{W2}$ | $Partly\ cloudy_{D-1}$ (binary) |
| T15 | 3 p.m. (binary) | $D1_{W3}$ | $Mostly\ cloudy_{D-1}$ (binary) |
| T16 | 4 p.m. (binary) | $D1_{W4}$ | $Cloudy_{D-1}$ (binary) |
| T17 | 5 p.m. (binary) | $D1_{Temp}$ | $Temperature_{D-1}$ (continuous) |
| T18 | 6 p.m. (binary) | $D1_{Humi}$ | $Humidity_{D-1}$ (continuous) |
| W1 | Clear (binary) | $D1_{WS}$ | $Wind\ speed_{D-1}$ (continuous) |
| W2 | Partly cloudy (binary) | $D1_{Solar}$ | $Solar\ irradiance_{D-1}$ (continuous) |

In this study, we addressed the issue of needing more data in the meteorological information collected. Approximately 0.1% of the total data for each category, including temperature, humidity, wind speed, and solar irradiation, needed to be included, with missing values indicated as −1. We employed linear interpolation to estimate these missing values given their continuous data characteristics. We utilized logistic regression to approximate missing values based on similarity with adjacent data for sky condition data, which were presented as categorical values from 1 to 4. To effectively reflect the periodicity of the date and maintain consistency with our previous study [7], we performed a day-ahead hourly solar irradiance forecasting using the same independent variables as shown in Table 1. We first converted the date information to Julian dates, ranging from 1 to 365 (common years) or 366 (leap years), where January 1 corresponds to 1 and December 31 to 365, or 366 leap years. Subsequently, we augmented the one-dimensional date data with continuous data in two-dimensional space using Equations (1) and (2) [6], [7]. This preprocessing of the date data, as illustrated in Figure 2, allowed for a more accurate and efficient representation of the periodicity within the collected meteorological data.

The sky condition comprises four categories on an interval scale from 1 to 4: clear, partly cloudy, mostly cloudy, and cloudy, with the cloud amount represented by eleven scales according to the climatology 1/10 method used by the KMA. To effectively represent these categorical data, we employed one-hot encoding, assigning a value of 1 to the binary variable for a specific sky condition and 0 for the other conditions. Similarly, we used one-hot encoding to represent the hour factor on an interval scale from 8 (8 a.m.) to 18 (6 p.m.), considering that solar irradiance typically peaks between 12 p.m. and 2 p.m. To account for recent trends in solar irradiance, we incorporated historical weather conditions from one day before the prediction time, including sky condition, temperature, humidity, wind speed, and solar irradiance, as independent variables. Consequently, our hybrid DL model construction dataset comprised 30 independent variables and solar irradiance as the

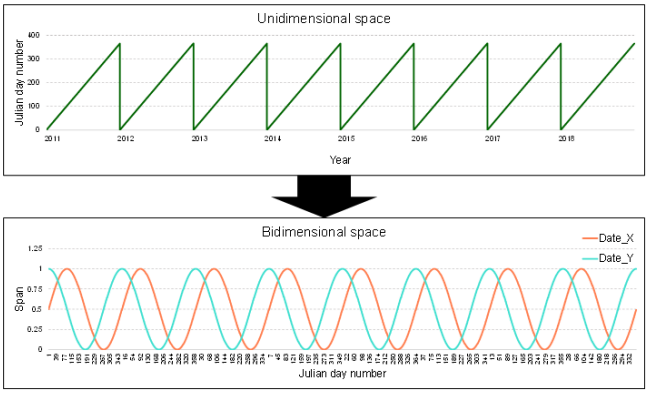



**Figure 2.** Visualization of Julian date data preprocessing.

$$Date_x = \sin(360° \times Julian\ Date/365 \text{ or } 366) \quad (1)$$

$$Date_y = \cos(360° \times Julian\ Date/365 \text{ or } 366) \quad (2)$$

dependent variable. This comprehensive representation of sky conditions, time intervals, and recent weather trends allowed for more accurate and adaptable predictions of intermittent solar irradiance in our model.

## III. Forecasting Model Construction

We aimed to develop a novel and precise hybrid DL model for short-term solar irradiance predictions in Jeju Island, South Korea. Our model combined the strengths of GRUs and TCN to effectively capture nonlinear weather parameters and address long-term dependencies in the data. GRUs, a type of recurrent neural network, are more computationally efficient than LSTM models and can handle long-term dependencies effectively [8]. On the other hand, TCN uses convolutional operations to solve long-term dependencies, resulting in high performance on long sequence data [9].

Our hybrid DL model consisted of an encoder and a decoder. The encoder was composed of three GRU layers with unit counts of 64, 32, and 16, and the final unit's state served as the input for the TCN layer. Convolution operations were performed with three kernels, and the final output was obtained through GlobalAveragePooling1D. The decoder part was configured with the same GRU layer as the encoder, and self-attention was applied at the output of the final decoder to weigh the importance of each time step in the prediction process [10].

The activation function is a critical component of the model that determines each neuron's output and can significantly impact the model's performance [6]. Therefore, we evaluated the performance of three activation functions: ReLU, SELU, and Leaky ReLU. ReLU maps negative values to 0 and positive values to their original value but can lead to the "dying ReLU" problem, where some neurons cannot learn. SELU has a non-zero mean and fixed variance, which helps maintain the activation values within a certain range, leading to faster convergence and improved performance compared to ReLU. Leaky ReLU is a variant of ReLU that allows a small positive slope for negative values, solving the "dying ReLU" problem. However, it can also suffer from the vanishing/exploding gradient problem.

In conclusion, our study utilized a hybrid DL model that combined GRU, TCN, and self-attention mechanisms to accurately predict solar irradiance for the next 11 time points in Jeju Island, South Korea. We also considered the impact of different activation functions on the model's performance.

## IV. Results and Discussion

The study used two solar irradiance datasets collected from Jeju Island regions, Ildo-1 dong and Gosan-ri, from 2011 to 2018. These datasets were divided into a training set (in-sample) and a test set (out-of-sample). The training set was used to train the DL models, while the test set was used to evaluate their prediction performance. The datasets were split into training and test sets at a ratio of 75:25, respectively. The training set contained data from 2011 to 2016, while the test set contained data from 2017 to 2018. This division ensured that the models were trained on sufficient data while being tested on unseen data to evaluate their generalization performance.

To evaluate the prediction performance of the models, the study used two metrics calculated using Equations (3)–(4): mean absolute error (MAE) and root mean square error (RMSE).

$$MAE = 1/n \times \sum|F_t - A_t|, \quad (3)$$

$$RMSE = (\sqrt{}(\sum (F_t - A_t)^2)/n, \quad (4)$$

where $A_t$ and $F_t$ represent the actual and forecasted values at time $t$, and $n$ indicates the number of observations. These metrics quantify the difference between the predicted and actual solar irradiance values. The MAE measures the average absolute difference between the predicted and actual values. In contrast, the RMSE measures the square root of the average squared difference between the predicted and actual values.

The proposed hybrid DL prediction model was implemented using Python version 3.8, with scikit-learn version 1.2.1, Tensorflow version 2.9.0, and Keras version 2.9.0. The random seed was fixed to 42 for reproducibility, and a batch size of 11 was used to train the model on 11 samples simultaneously in each batch. The model was trained for 50 epochs, and the hyperparameter settings were chosen through trial and error and best practices in the DL literature. These settings allowed the model to achieve good prediction performance while avoiding overfitting and ensuring the reproducibility of the results.

Table 2 compares the performance of a hybrid DL model in predicting solar irradiance using different activation functions (i.e., ReLU, SELU, and Leaky ReLU) in two regions of South Korea. The results show that the SELU function had the lowest average RMSE and MAE values for both regions, while Leaky ReLU and ReLU performed worse in one region each. This suggests that the choice of activation function can significantly affect model performance for solar irradiance prediction. SELU may be a better choice due to its self-normalizing and efficient learning properties.

SELU is a variant of ReLU with a self-normalizing property that helps mitigate the vanishing gradient problem in deep neural networks [11]. This property ensures that the output of each layer in the network has a zero mean and unit variance, regardless of weight initialization. Additionally, SELU has a non-zero mean and fixed variance that helps maintain activation values within a certain range, leading to faster convergence and

TABLE 2.
PERFORMANCE COMPARISON BY ACTIVATION FUNCTIONS FOR ILDO-1 DONG AND GOSAN-RI (UNITS: MJ/M²).

| Steps | Ildo-1 dong | | | | | | Gosan-ri | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ReLU | | SELU | | Leaky ReLU | | ReLU | | SELU | | Leaky ReLU | |
| | *MAE* | *RMSE* | *MAE* | *RMSE* | *MAE* | *RMSE* | *MAE* | *RMSE* | *MAE* | *RMSE* | *MAE* | *RMSE* |
| 1 | 0.366 | 0.253 | 0.394 | 0.279 | 0.384 | 0.268 | 0.252 | 0.347 | 0.231 | 0.336 | 0.252 | 0.347 |
| 2 | 0.453 | 0.312 | 0.459 | 0.323 | 0.463 | 0.321 | 0.293 | 0.405 | 0.283 | 0.411 | 0.293 | 0.405 |
| 3 | 0.497 | 0.341 | 0.498 | 0.349 | 0.508 | 0.350 | 0.322 | 0.452 | 0.310 | 0.450 | 0.322 | 0.452 |
| 4 | 0.525 | 0.359 | 0.526 | 0.365 | 0.539 | 0.370 | 0.346 | 0.484 | 0.333 | 0.483 | 0.346 | 0.484 |
| 5 | 0.543 | 0.370 | 0.543 | 0.374 | 0.565 | 0.388 | 0.359 | 0.504 | 0.350 | 0.509 | 0.359 | 0.504 |
| 6 | 0.556 | 0.380 | 0.556 | 0.379 | 0.587 | 0.400 | 0.369 | 0.518 | 0.361 | 0.524 | 0.369 | 0.518 |
| 7 | 0.567 | 0.390 | 0.564 | 0.384 | 0.601 | 0.410 | 0.379 | 0.531 | 0.367 | 0.532 | 0.379 | 0.531 |
| 8 | 0.581 | 0.403 | 0.571 | 0.390 | 0.612 | 0.420 | 0.390 | 0.546 | 0.374 | 0.540 | 0.390 | 0.546 |
| 9 | 0.596 | 0.416 | 0.577 | 0.396 | 0.623 | 0.433 | 0.404 | 0.561 | 0.384 | 0.548 | 0.404 | 0.561 |
| 10 | 0.612 | 0.429 | 0.588 | 0.407 | 0.633 | 0.444 | 0.420 | 0.579 | 0.394 | 0.558 | 0.420 | 0.579 |
| 11 | 0.629 | 0.442 | 0.601 | 0.418 | 0.644 | 0.457 | 0.421 | 0.584 | 0.402 | 0.568 | 0.421 | 0.584 |
| Avg. | 0.539 | 0.372 | 0.534 | 0.369 | 0.560 | 0.387 | 0.360 | 0.501 | 0.344 | 0.496 | 0.360 | 0.501 |

better performance than ReLU. SELU outperformed other activation functions in benchmark datasets and DL models.

In predicting solar irradiance, SELU's ability to handle long-term dependencies and mitigate the vanishing gradient problem is particularly relevant due to the cyclical nature of solar irradiance data. Furthermore, SELU's self-normalizing property helps the model learn more effectively from the data, leading to better prediction performance. Therefore, SELU is a good choice for DL models, including those used for solar irradiance prediction.

## V. CONCLUSIONS

This paper proposes a hybrid DL model that accurately predicts solar irradiance by considering various meteorological variables and historical solar irradiance data. The model combines GRU and TCN to learn short-term and long-term patterns in time series data. At the same time, self-attention is applied to identify the importance of each variable in predicting solar irradiance. Results show that SELU is the most suitable activation function for solar irradiance prediction models, especially in areas with large deviations in observed solar irradiance values. These findings highlight the importance of solar irradiance prediction for renewable energy generation and demonstrate the potential of the proposed model for practical energy planning in real-world environments. Additionally, this study presents the possibility of applying digital twins to developing smart energy management systems, contributing to advancing related technologies.

Future research can focus on the following aspects. Firstly, the proposed hybrid DL model can be further optimized by exploring other activation functions and architectures that can better capture the nonlinear characteristics of solar irradiance data. Secondly, more meteorological variables can be included to improve the accuracy of solar irradiance prediction further. Thirdly, the proposed model can be applied to other regions with different climate conditions to evaluate its generalizability and scalability. Finally, future research can also explore integrating the proposed model with smart grid systems and energy storage technologies to achieve optimal energy management and reduce the impact of renewable energy intermittency.

## ACKNOWLEDGEMENTS

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2023-2020-0-01832) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019M3F2A1073179).

## REFERENCES

[1] C. Ghenai, L. A. Husein, M. Al Nahlawi, A. K. Hamid, and M. Bettayeb, "Recent trends of digital twin technologies in the energy sector: A comprehensive review," *Sustainable Energy Technologies and Assessments*, vol. 54, p. 102837, 2022.

[2] M. Ma, B. He, R. Shen, Y. Wang, and N. Wang, "An adaptive interval power forecasting method for photovoltaic plant and its optimization," *Sustainable Energy Technologies and Assessments*, vol. 52, p. 102360, 2022.

[3] A. Bhatt, W. Ongsakul, and J. G. Singh, "Sliding window approach with first-order differencing for very short-term solar irradiance forecasting using deep learning models," *Sustainable Energy Technologies and Assessments*, vol. 50, p. 101864, 2022.

[4] R. Chang, L. Bai, and C.-H. Hsu, "Solar power generation prediction based on deep learning," *Sustainable Energy Technologies and Assessments*, vol. 47, p. 101354, 2021.

[5] X. Luo, and D. Zhang, "An adaptive deep learning framework for day-ahead forecasting of photovoltaic power generation," *Sustainable Energy Technologies and Assessments*, vol. 52, p. 102326, 2022.

[6] J. Moon, S. Park, S. Rho, and E. Hwang, "A comparative analysis of artificial neural network architectures for building energy consumption forecasting," *International Journal of Distributed Sensor Networks*, vol. 15, no. 9, p. 1550147719877616, 2019.

[7] J. Park, J. Moon, S. Jung, and E. Hwang, "Multistep-ahead solar radiation forecasting scheme based on the light gradient boosting machine: A case study of Jeju Island," *Remote Sensing*, vol. 12, no. 14, p. 2271, 2020.

[8] M. Xia, H. Shao, X. Ma, and C. W. de Silva, "A stacked GRU-RNN-based approach for predicting renewable energy and electricity load for smart grid operation," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 10, pp. 7050–7059, 2021.

[9] P. Hewage, A. Behera, M. Trovati, E. Pereira, M. Ghahremani, F. Palmieri, and Y. Liu, "Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station," *Soft Computing*, vol. 24, pp. 16453–16482, 2020.

[10] F. Yu, L. Wang, Q. Jiang, Q. Yan, and S. Qiao, "Self-Attention-Based Short-Term Load Forecasting Considering Demand-Side Management," *Energies*, vol. 15, no. 12, p. 4198, 2022.

[11] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-normalizing neural networks," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

# Sleep Quality Monitoring System Based On CSHQ using pressure sensor

Sang-Hun Lee[1*], Byeong-Gwon Kang[2] and Jung-Yeon Kim[3]
[1]Dept. of ICT Convergence, Soonchunhyang University, Asan, South Korea
[2]Dept. of Information and Communication Engineering, Soonchunhyang University, Asan, South Korea
[3]ICT Convergence Research Center, Soonchunhyang University, Asan, South Korea
*Contact: tkdgns9730@sch.ac.kr, phone +82-10-2584-9730

*Abstract*—**Sleep plays a critical role in our overall health and well-being, accounting for about one-third of our lives. While studies have shown that sleep aids in physical and mental recovery, and enhances immune system and cognitive function, insufficient sleep can be detrimental to our health. Despite the growing awareness of the importance of sleep, the number of patients with sleep disorders, such as insomnia, continues to rise. To address this issue, many self-diagnostic checklists for sleep disorders have been developed. However, relying on participants' memory only can result in inaccurate results. In this paper, we propose a sensor-based approach to overcome the limitations of using checklists for gathering information on a specific target. We use an 11x11 array of pressure sensors to receive real-time pressure values, which are preprocessed and used to calculate moving averages and standard deviations. Based on these values, a control limit is imposed, and pressure values that exceed this limit indicate the start of movement. A numerical checklist is then generated from these values, offering more precise and reliable sleep disorder diagnosis results.**

## I. INTRODUCTION

Sleep plays an important role in human health and daily life. However, research shows that children with developmental disabilities are more likely to have a negative impact on their sleep. This is because there may be relative deficiencies in self-awareness and sleep regulation. Therefore, systematic sleep evaluation is needed to measure and solve sleep problems in children with developmental disabilities [1], [2].

There are various indicators to evaluate sleep, but they often rely on parents' perceptions, which can include subjective factors and limit the accuracy of the evaluation [3].

Recently, various sleep evaluation technologies have been developed by combining sleep-related indicators with various sensors to overcome the aforementioned limitations.

Among them, we evaluate sleep based on Children's Sleep Habits Questionnaire (CSHQ) and pressure sensors, which are indicators focused on evaluating sleep problems in children with developmental disabilities [4].

The reason for using pressure sensors is that, unlike other sensors used for sleep measurement, they require direct contact with the bed or mattress, allowing for more detailed information to be collected. Therefore, using pressure sensors for sleep measurement ensures high accuracy in measuring and analyzing various physiological signals that occur during sleep, such as movements, breathing, and heart rate.

However, using pressure sensors also has its drawbacks, as the user's freedom of movement may be limited because they need to stay in contact with the pressure sensors on some specified locations of mattress [5].

Therefore, in this study, in order to solve the problem of sleep indicators with limited accuracy, more accurate sleep evaluation is performed using CSHQ and pressure sensors.

## II. TOOLS AND EQUIPMENT

In this paper, we quantified the items that can be observed in the CSHQ using an 11x11 array of pressure sensors, and we plan to analyze the items related with movements during sleepy time.

### *A. Children's Sleep Habits Questionnaire (CSHQ)*

The Children's Sleep Habits Questionnaire (CSHQ) is a validated tool used to assess the sleep behaviors of children between the ages of 4 and 12. Parents complete a survey consisting of 52 questions to evaluate the frequency of their child's sleep behavior patterns. The questionnaire has excellent reliability and validity and takes approximately 10 minutes to complete. By identifying sleep disorders such as bedtime resistance and night waking, the CSHQ allows for early intervention and treatment to improve children's emotional health, academic performance, and relationships.

We quantified the items in CSHQ that can be assessed with pressure sensors and classified those related to movement, presenting them in Table II

TABLE I
The items to be analysed in the quantified CSHQ

| division | Check(O,X) |
|---|---|
| Child struggles in bed at bedtime | |
| The child is unable to move or sit still while sleeping | |

### *B. Pressure sensor mat*

A fiber-based metro material was selected for a floor mat electrode to monitor foot voltage, and a printing process was used to form the electrode. An FSR-type pressure sensor was

chosen for monitoring bed behavior, and the electrode configuration was improved to increase sensitivity.

A pressure sensor was also constructed for a mattress cover using a fiber assembly process with a conductor for durability and stability. A 121-channel pressure sensor module was developed and its wireless data transmission and reception performance was evaluated.

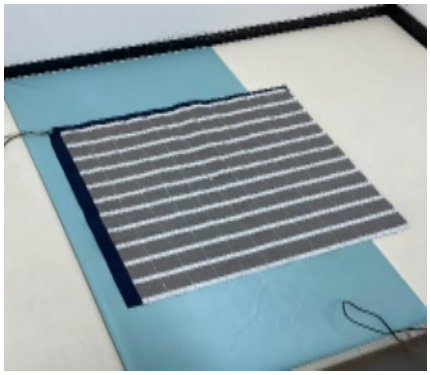


Fig. 1 pressure sensor mat

## III. EXPERERIMENT PROTOCOL

As for the experimental protocol, 5 experiments were conducted from healthy young adults aged 20-30 years, and tossing and turning 1 time (test1), tossing and turning 2 times (test2), tossing and turning 3 times (test3), and tossing 4 times (test4), and tossing 5 times while lying on a mat. The number (test 5) was conducted a total of 5 times for 30 seconds each.


Fig. 2 experimental protocol

### A. Pre-processing

To improve the quality of the pressure values obtained from the pressure sensor mat and remove noise from the data, a symmetric Gaussian low-pass filter of size 10 with a standard deviation of 0.4 was applied. This minimized the impact of noise on the pressure values.

As shown in Figure 3, samples of the original data and the preprocessed data are displayed. It can be seen that noise and unnecessary high-frequency components have been reduced, and the quality of the data has been improved.

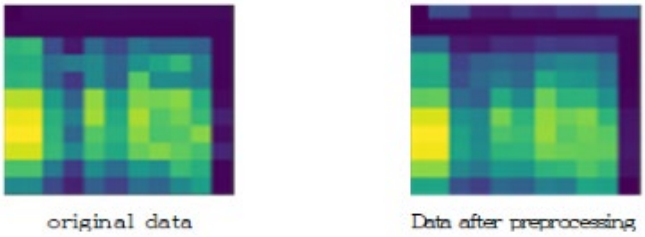



Fig. 3 Original data and pre-processed data

### B. Movement identification

There have been many studies using pressure sensors to detect movement, and using a controlled threshold value has proven to be the most accurate method. Therefore, the following method was used to detect movement [6].

The first step in motion detection is to calculate the expected value of the data and determine the moving average and standard deviation of the moving average. The formulas for the moving average and standard deviation of the moving average are as follows.

$$\mathrm{MA(n)} = \left(\frac{1}{n}\right) \times (x_1 + x_2 + x_3 \ldots + x_n) \qquad (1)$$

$$\mathrm{SD(n)} = \sqrt{\frac{1}{n} \times \sum_{i=0}^{n} (x_i - MA(n))^2} \qquad (2)$$

A control limit is imposed based on the moving average and standard deviation of the moving average, and any pressure above the imposed control limit indicates the onset of motion.

Therefore, the upper and lower control limits are set as follows.

$$\mathrm{UCL} = \mathrm{MA(n)} + \left(3 \times \mathrm{SD(n)}\right) \qquad (3)$$

$$\mathrm{LCL} = \mathrm{MA(n)} - \left(3 \times \mathrm{SD(n)}\right) \qquad (4)$$

Controlled threshold was used to detect motion, and an upper control limit (UCL) and a lower control limit (LCL) were established based on the moving average and standard deviation of the moving average.

Moving average is commonly used in time series data to identify regular patterns by calculating the average of the data over a given period. The standard deviation of the moving average represents the standard deviation of the deviations from the moving average.

Equation (1) is the formula for calculating the moving average. 'n' represents the number of data points during a given period. It is the sum of each data value divided by 'n', and the moving average represents the average of the data.

Equation (2) is the formula for calculating the standard deviation of the moving average. The sigma symbol represents the sum of each data value, and 'i' represents the index of the data value.

By squaring the difference between the moving average and each data value, dividing by 'n', and taking the square root, the standard deviation of the moving average can be calculated.

Equations (3) and (4) are the formulas for calculating the upper and lower control limits, respectively.

The standard deviation of the moving average is multiplied by 3 and added to or subtracted from the moving average.

These control lines are used as criteria for determining whether motion has occurred. If the data value exceeds the upper control limit or falls below the lower control limit, this is considered to be motion detected.

## IV. Experiment Result

Figures 4 and 5 show samples of the pressure values obtained from the pressure sensor mat with movements identified based on the controlled threshold method for each test.

Sometimes, as shown in Figure 4, the number of movements and movement detection occurrences may be the same.

However, since the experiment is conducted on human subjects, there may be cases where movements are slow or not precisely controlled, resulting in undetected movements or cases where movements are detected twice. Figure 5 depicts a scenario where movement is detected twice.

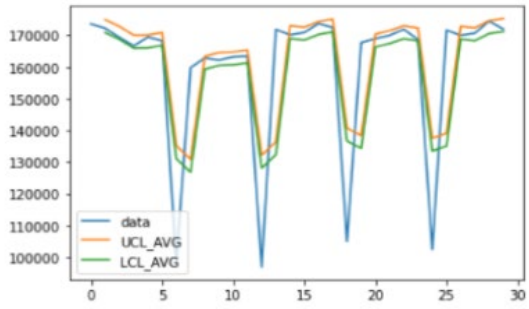



Fig. 4 Accurate motion detection sample

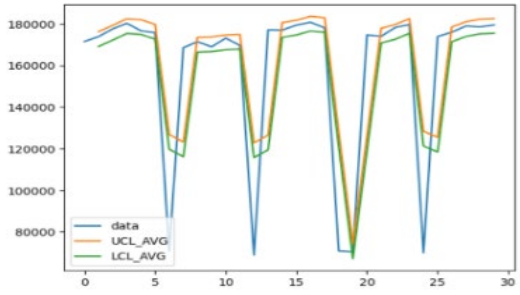



Fig. 5 Samples with incorrect motion detection

The tables below represent the recognition rates for each test. For example, in Test 1 (where there is one movement), if the exact movement is detected in all 5 trials, the recognition rate would be 100%. If it is detected in 4 out of 5 trials, the rate would be 80%, and so on, with 60% for 3 times, 40% for 2 times, and 20% for 1 time.

Table 2 shows the recognition rate when the experiment was conducted without control over movement, and the rate was 68%. Table 3 represents the experiment with control over movement, resulting in a recognition rate of 88%, which is 20% higher than the rate achieved without control over movement. As shown by the experiments above, the recognition rate varies depending on the control over movement.

TABLE Ⅱ

Recognition rate when experimenting without controlling motion

| TEST | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| accuracy | 80% | 60% | 40% | 80% | 80% |
| 68% | | | | | |

TABLE Ⅲ

Recognition rate in motion control experiment

| TEST | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| accuracy | 100% | 100% | 60% | 80% | 100% |
| 88% | | | | | |

In summary, our results demonstrate the effectiveness of us ing controlled threshold values to detect movements based on t he pressure values obtained from pressure sensor mats. Howev er, there may be overlapping motion detections in cases where the movement is slow.

Further research should be conducted to address this issue.

## V. Conclusions

This paper presents a system for experimental results that shows that motion can be effectively detected using pressure values obtained from a pressure sensor mat with a controlled threshold based on the moving average and standard deviation of the moving average. The movements identified were in good agreement with the movements expected based on the experimental protocol, indicating the accuracy and reliability of this method.

Based on the results presented in this paper, further research can be conducted to validate the effectiveness of this system in real-world scenarios, such as in home settings where individuals have different sleeping patterns and behaviors.

Additionally, future studies can explore the potential of using this system for detecting other types of movements, such as respiratory movements, and how it can be integrated with other sleep monitoring technologies to provide a more comprehensive sleep analysis.

We anticipate that this system can be used to more objectively quantify sleep-related indicators.

## Acknowledgment

"This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) program(IITP-2023-2020-0-01832) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation)"

## REFERENCES

[1] Malow, B. A., Crowe, C., Henderson, L., McGrew, S. G., & Wang, L. (2009). A sleep habits questionnaire for children with autism spectrum disorders. Journal of Child Neurology, 24(8), 979-985.

[2] Giannotti, F., Cortesi, F., Cerquiglini, A., Miraglia, D., Vagnoni, C., & Sebastiani, T. (2008). An investigation of sleep characteristics, EEG abnormalities and epilepsy in developmentally regressed and non-regressed children with autism. Journal of Autism and Developmental Disorders, 38(10), 1888-1897.

[3] Mindell, J. A., & Owens, J. A. (2015). A Clinical Guide to Pediatric Sleep: Diagnosis and Management of Sleep Problems. Lippincott Williams & Wilkins.

[4] Owens, J. A., Spirito, A., & McGuinn, M. (2000). The Children's Sleep Habits Questionnaire (CSHQ): psychometric properties of a survey instrument for school-aged children. Sleep, 23(8), 1043-1051.

[5] "Pressure Sensitive Bed: A Novel Method for Monitoring Sleep Apnea" by Lim et al. (2005)
[6] Gilakjani, S. Soleimani, et al. "Movement detection with adaptive window length for unobtrusive bed-based pressure-sensor array." 2017 IEEE International Symposium on Medical Measurements and Applications (MeMeA). IEEE, 2017

# A Novel Adaptive Sliding Mode Control of Robot Manipulator with Time Delays for Assisted Living

Hyuk Mo An[1] and Seok Young Lee[1,*]
[1]*Department of ICT Convergence Engineering, Soonchunhyang University, Asan, South Korea*
*Contact: suk122@sch.ac.kr

*Abstract*—**Robotic manipulators for assisted living to be used in IoT environments, need to be controllable even in environments with time delays. This paper proposes an adaptive sliding mode controller for a manipulator system with time delays and uncertain disturbances. The robustness of the controller against disturbances and time delays is verified through the Lyapunov stability theorem. Simulation results demonstrate the effectiveness of the proposed control method.**

## I. INTRODUCTION

Robot manipulators have been utilized in a variety of environments such as factories and homes. Due to such flexibility, these types of robots can be used for assisted living in combination with IoT. It must be robust against time delays in dynamic systems to provide stable assisted living in a network environment. In addition, the manipulator needs robust control algorithms to enable the desired action even in the event of an impact or unexpected situation.

Robust control in the presence of disturbances has been a long-standing interest of the academic community in the field of systems [1]. To achieve efficient control, sliding mode control has been a popular technique due to its robustness against disturbances in system dynamics. Since sliding mode control forces the system state variables onto a sliding surface, regardless of disturbance, system states converge in a finite time along the surface. It can be proven through the Lyapunov stability theorem and has been widely used in manipulator control systems [2].

The main issue with sliding mode control is a chattering phenomenon and the need to know the upper bound of uncertainties. If the control gain value is greater than the upper bound of the uncertainties, the system is robust to disturbances and uncertainty. In situations where there is no information about the upper bound of the uncertainty, the problem can be solved by using the adaptive sliding mode control method proposed in [3].

Sliding mode control has been applied to systems with time delays in [4]. Since existing studies adjust system parameters based on LMI, which requires a lot of computation. It is very difficult to apply the LMI-based approaches to real-time robotic systems. To reduce the computation, this paper proposes a novel sliding mode controller based on signum function.

The mechanisms presented in this paper are effective in time-delayed environments with an unknown upper bound of disturbances. We applied sliding mode control using adaptive gains to a 3-DOF manipulator system with a time delay. Here, we combine the saturation function and the signum function to guarantee stability even with time delays. Compared to the previous results, our method shows better performance in terms of stability.

## II. PRELIMINARIES

### A. Manipulator system

Consider the n-DOF manipulator dynamics consisting of the following formula.

$$M(q)\ddot{q} + C(q,\dot{q}) + G(q) = u(t) + d, \tag{1}$$

where system parameters $q, \dot{q}, \ddot{q} \in \mathbb{R}^n$ are position, velocity, and acceleration vectors of each manipulator joint, respectively. $M(q) \in \mathbb{R}^{n\times n}$ is the inertia matrix, which is positive definite. $C(q) \in \mathbb{R}^n$ is the centripetal and Coriolis vector. $G(q) \in \mathbb{R}^n$ is the gravitation vector. The input torque $u(t) \in \mathbb{R}^n$ is the force applied to each joint of the manipulator. $d \in \mathbb{R}^n$ is the unknown disturbance torque of the manipulator system. If we organize (1) with respect to $\ddot{q}$, we get the formula below

$$\ddot{q} = M^{-1}\big(u(t) - C(q,\dot{q}) - G(q)\big). \tag{2}$$

Next, we define the system error below

$$e = q_d - q, \tag{3}$$

where $q_d \in \mathbb{R}^n$ denotes the desired position trajectory. Based on the error dynamics (3) defined above, $\dot{e}$, $\ddot{e}$ stand for velocity error, and acceleration error, respectively.

### B. Adaptive sliding mode control

This section applies the adaptive sliding mode control proposed in [3] to the n-DOF manipulator dynamics (1). We define the following sliding surface,

$$s = \dot{e} + \Lambda e. \tag{4}$$

where $\Lambda \in \mathbb{R}^{n\times n}$ denotes the positive diagonal matrix. Utilizing the system dynamics (2), the derivative of the sliding surface can be obtained as follows,

$$\begin{aligned}\dot{s} &= \ddot{e} + \Lambda\,\dot{e} = \ddot{q} - \ddot{q}_d + \Lambda\,\dot{e} \\ &= \ddot{q}_d - M^{-1}(u - C - G) + \Lambda\,\dot{e}.\end{aligned} \tag{5}$$

The paper [3], develops $u$ by splitting it into two terms

$$\begin{aligned}u &= u_0 + u_1 \\ u_0 &= C + G + M(\ddot{q}_d + \Lambda\,\dot{e}) \\ u_1 &= M\widehat{K}\,sgn(s).\end{aligned} \tag{6}$$

The control gain $\hat{K} \in \mathbb{R}^{n\times n}$ is a diagonal gain matrix. $K^* \in \mathbb{R}^{n\times n}$ is the gain at the point where it stops increasing. $\alpha \in \mathbb{R}$ is a positive adaptive gain. The adaptive law is defined as the formula below

$$\dot{\hat{K}}(i,i) = \alpha|s_i|. \tag{7}$$

$\hat{K}$ increases until the sliding surface reaches zero.

## III. Main Result

The system control scheme presented in Section 2 has not been validated for systems with time delays. To apply it to systems with time delays, we propose a new control input.

### A. A novel adaptive sliding mode for time delay systems

When the system input has a fixed time delay $\tau$, we propose the controller shown below

$$\begin{aligned} u_0 &= C + G + M(\ddot{q}_d + \Lambda\,\dot{e}) \\ u_1 &= M\{\frac{1}{2}\hat{K}(t-\tau)sgn(s(t-\tau)) \\ &\quad + \frac{1}{2}\left(\hat{K}(t-\tau)+\epsilon\right)sat\left(s(t-\tau)\right)\}. \end{aligned} \tag{8}$$

$\epsilon \in \mathbb{R}$ is small positive constant. Saturation function $sat(s)$ denotes

$$sat(s) = \begin{cases} s/\sigma, & |s| \le \sigma \\ sgn(s), & |s| > \sigma \end{cases} \tag{9}$$

### B. Stability analysis

The stability of the proposed controller is verified through the Lyapunov stability theorem. Consider the following Lyapunov function

$$V = \frac{1}{2}s(t)^T s(t) + \frac{1}{2}\alpha\left(\hat{K}-K^*\right)^2 \ge 0. \tag{10}$$

The derivate of Lyapunov function V can be obtained as follows

$$\begin{aligned} \dot{V} &= \dot{s}s + \alpha\left(\hat{K}-K^*\right)\dot{\hat{K}} \\ &= s(t-\tau)\{\frac{1}{2}\left(-\hat{K}(t-\tau)\right)sgn\left(s(t-\tau)\right) \\ &\quad -\frac{1}{2}(\hat{K}(t-\tau)+\epsilon)sat(s(t-\tau))\} \\ &\quad +\alpha^2\left(\hat{K}-K^*\right)|s| \end{aligned} \tag{11}$$

Compared to the signum function, a gradient of the saturation function is more smaller. Thus we assume that the functions below are similar.

$$sat(s(t-\tau)) \approx sat(s) \tag{12}$$

Then, if $s \ge 0, \dot{V} \approx s(-\epsilon\, sgn(t-\tau)) + \alpha^2\left(\hat{K}-K^*\right)|s| < 0$. If $s < 0, \dot{V} \approx s\left(\epsilon\, sgn(t-\tau)\right) + \alpha^2\left(\hat{K}-K^*\right)|s| < 0$.

The reason for using the saturation function instead of the signum function is to smooth out the discontinuous difference between $sgn(s(t-\tau))$ and $sgn\left(s(t)\right)$ caused by the time delay.

## IV. Simulation Result

The proposed controller is simulated with the presence of a time delay in the 3-DOF manipulator dynamics. In the simulation, we set desired trajectory $q_d = [q_{1d}, q_{2d}, q_{3d}]^T$, $q_{1d} = \frac{\pi}{6}(1-\cos(1.5\pi t)+\sin(\pi t))$, $q_{2d} = \frac{\pi}{6}(1-\cos(2\pi t)+\sin(1.5\pi t))$, and $q_{3d} = \frac{\pi}{6}(1-\cos(3\pi t)+\sin(0.7\pi t))$. The disturbance $d = [2\cos(2t), 2\sin(2t), 2\cos(3t)]^T$, adaptive gain $\alpha = 5$, $\Lambda = \mathrm{diag}[13,13,13]$, and boundary layer $\sigma = 0.2$. The manipulator length is $l_1 = 0.5\mathrm{m}, l_2 = 0.5\mathrm{m}, l_3 = 0.5\mathrm{m}$ and mass is $m_1 = 0.8\mathrm{kg}, m_2 = 0.4\mathrm{kg}, m_3 = 0.4\mathrm{kg}$. The system input delay is defined by $\tau = 0.0001$.

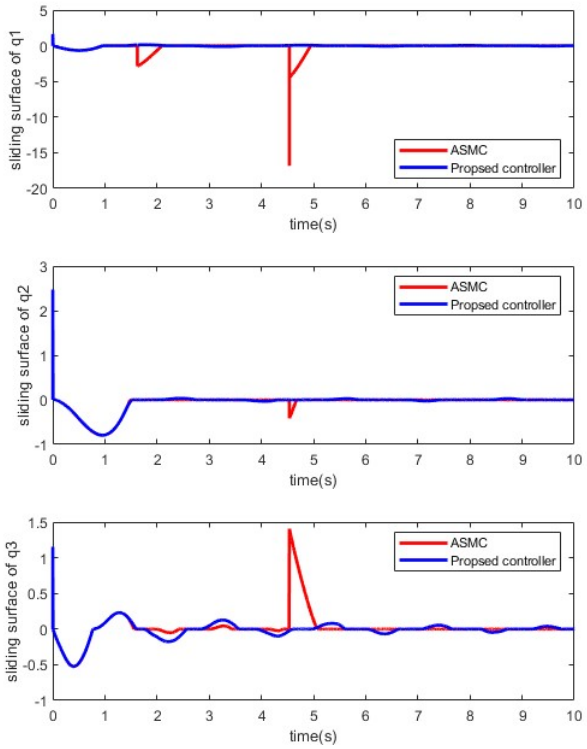



Fig. 1 Comparisons of the sliding surface in each joint of the manipulator.

The sliding surface of the proposed controller and the adaptive sliding mode controller for the 3-DOF manipulator is shown in Fig. 1. It can be seen that the proposed controller reduces the rapid changes in the sliding surface compared to the adaptive sliding mode controller [3].

## V. Conclusions

In this paper, we proposed a novel adaptive sliding mode control for a time-delay system. The proposed controller uses the saturation function as a replacement for $sgn\left(s(t)\right)$. It was confirmed that the proposed controller can stabilize the system even when the upper bound of the disturbance is unknown, and the system can be controlled even in the presence of time delay.

## Acknowledgment

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) program(IITP-2023-2020-0-01832) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation)

## References

[1] V.I. Utkin, Sliding Modes in Control and Optimization. New York: Springer-Verlag, 1992.

[2] J. Baek, M. Jin, S. Han. Sattler, 'A new adaptive sliding-mode control scheme for application to robot manipulators' IEEE Transactions on Industrial Electronics. 63 (6) 3628-3637 (2016)
[3] Huang, Y.J, Kuo, T.C, and Chang, S.H Adaptive sliding mode control for nonlinear systems with uncertain parameter' IEEE Transactions on System, Man, and Cybernetics – Part B: Cybernetics 38, 534-539 (2008)
[4] Yuanqing Xia and Yingmin Jia. "Robust Sliding-Mode Control for Uncertain Time-Delay Systems: An Lmi Approach." IEEE Transactions on Automatic Control 48 (6): 1086–92(2003)

# MEASURE: Multimodal Exercise Assistance System Using Remote Equipment

Daniel Poul Mtowe, Jin-Hyuck Park, and Dong Min Kim*
*Soonchunhyang University, Asan, 31538, Korea*
*Contact: dmk@sch.ac.kr, phone +82 41 530 1535

*Abstract*— **The increasing number of elderly individuals in need of cognitive rehabilitation has prompted a need for further research on effective treatment methods. However, current limitations in access to professional therapy services have hindered progress in this area. To address this gap, we propose MEASURE: a multimodal exercise assistance system using remote equipment to provide real-time monitoring, evaluation, and feedback to enhance users' physical and cognitive abilities. In this system, therapists create customized exercise programs tailored to individual needs and goals in the form of virtual games connected through client-server architecture on the internet. Patients choose exercises presented as virtual games, allowing for control and interaction with virtual objects that require cognitive ability to follow instructions and complete tasks. The proposed solution is a user-friendly software that leverages advancements in technology for physical rehabilitation and multimodal interaction. MEASURE offers a range of features that support cognitive rehabilitation and enhance the effectiveness of exercise programs for elderly individuals.**

## I. Introduction

Physical therapy (PT) is a fundamental form of rehabilitation that plays a critical role in helping patients restore and improve their physical abilities. Given the diverse needs of patients, a range of therapeutic techniques may be required to address specific impairments and facilitate the recovery of optimal physical strength and cognitive function. Consistent engagement in PT activities over an extended period of time is necessary for patients to achieve a full return to normal daily activities. Patients need to continue PT activities at home after initial evaluation and training with therapists at a rehabilitation center. This is due to the inability to train multiple patients with a limited number of therapists, as well as the high direct and indirect costs to patients. Another problem arises when patients do not perform exercises at home on a regular and consistent basis, resulting in delays in achieving desired rehabilitation outcomes. On the other hand, it would be inconvenient and expensive for ordinary patients to visit the clinic frequently. In addition, the clinic is unable to dedicate additional resources to non-critical patients. If patients are able to connect with therapists and perform exercises at home, remote exercise management could be a potential solution to these difficulties. This can reduce the frequency of visits, lower costs, prevent delays in training critically ill patients, and eliminate the problem of limited patient training sessions. It will therefore provide countless benefits to all stakeholders, including patients, therapists, families and physicians.

Our proposed system aims to assist patients with their exercises in a pleasant and harmonious environment without requiring the presence of an expensive therapist. The therapist's instructions for the patient's tasks are pre-configured in the system for evaluation and monitoring purposes. The proposed system monitors the recovery process and provides further guidance as needed. The correctness of the tasks performed by the patient influences the healing process.

We put our MEASURE concepts into practice in the prototype system shown in Figure 1, which consists of three main parts: a camera, a screen, and a computing unit. The camera captures all the interaction activities of the patient's rehabilitation task and sends them to the computing unit for processing and analysis. The computing unit monitors and evaluates the interaction between the patient and the selected game by comparing the instructions to be followed and the thresholds to be reached with the patient's actions. Finally, the analysis reports are uploaded to the user's screen as feedback.

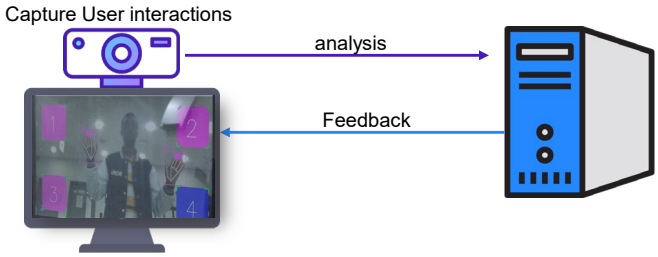



Fig. 1 System Block Diagram

In summary, we propose a multimodal exercise assistance system that uses remote devices to detect the user's performance and provide individualized and comprehensive guidance to improve the user's physical and cognitive abilities. Our proposed system uses image analysis to detect hand movements, monitor interactions between the user's hands and the virtual tasks they interact with, and assess the user's performance. We have developed a web-based in-home rehabilitation system that allows patients to complete rehabilitation at home, providing therapy to users who would otherwise be unable to receive it. Our primary goal is to provide a solution that complements existing rehabilitation training techniques and procedures, benefiting both patients and therapists and enabling effective

rehabilitation programs. The goal is to increase the quantity and quality of physical therapy sessions and to promote remote collaborative rehabilitation between patients and therapists. Our work focuses primarily on the design and architecture of MEASURE. Unlike other solutions, ours is non-invasive and relies solely on visual algorithms.

## II. Related Work

Over the past decade, significant advances have been made in the field of physical rehabilitation. One of the primary concerns of patients undergoing rehabilitation is to regain their ability to walk, which involves a series of exercises that can rapidly improve their range of motion and facilitate daily activities [1-3]. Despite the increasing popularity of clinical rehabilitation, there are growing concerns about its cost and effectiveness. Typically, long-term and intensive exercise programs are required for full recovery, which is time-consuming, expensive, and challenging. To address these issues, some researchers have explored alternative approaches such as video recording and virtual reality technology to monitor patients' movements and improve efficiency [4]. In addition, well-designed video games, such as motion-controlled games, have been found to be a valuable adjunct to standard physical therapy to increase patient participation and motivation [5]. These studies highlight the need for innovative and cost-effective approaches to rehabilitation that can improve patient outcomes and increase the efficiency of clinical rehabilitation programs.

Human pose estimation has been a widely researched area for several decades. Before the inception of convolution neural networks (CNNs), researchers developed various techniques to detect the main joint sites in images [6] [7]. Since 2012, CNNs have become the dominant approach in human pose estimation research [8-10]. For example, Cascade Pyramid Networks [10] have been proposed to integrate feature information from multiple scale representation maps, while HRNet [11] maintains a high-resolution feature representation branch throughout the architecture.

## III. Multimodal Exercise Assistance System Using Remote Equipment

Our system monitors and evaluates interactions between patients and customized exercise programs tailored to their individual needs. The exercises are presented to the patient in the form of virtual games connected through a client-server architecture, with the goal of improving the user's physical and cognitive abilities. MEASURE can identify and track the movement of the patient's hands, captured in live video files, as they perform the exercises, without the need for any wearable accessories on the human body. The exercises in our system are designed by therapists and tailored to each individual's specific needs and goals. Given the goals and instructions for each exercise provided by the therapist, our system evaluates and provides real-time feedback to the patient as they perform the exercise to encourage them to achieve their intended goals. With this level of monitoring and feedback, our system ensures rapid recovery. The proposed solution uses user-friendly software that takes advantage of advances in physical rehabilitation technology and multimodal interaction. The overall system architecture of MEASURE is shown in Figure 2 below.

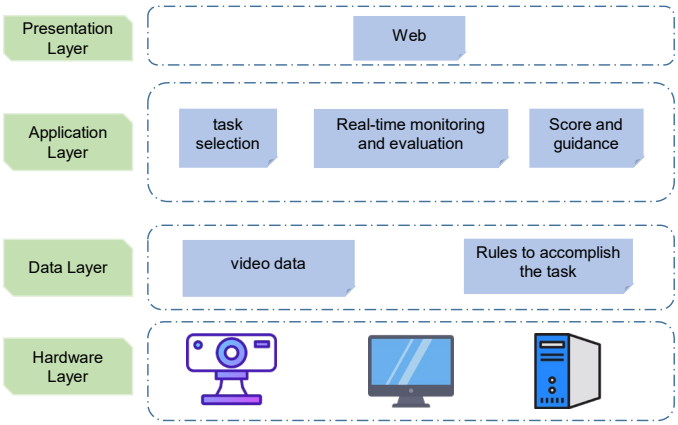



Fig. 2 System Architecture

Our hand tracking solution uses a machine learning (ML) pipeline consisting of two models that work together to achieve the desired result. This pipeline includes a hand detector that processes a full input image and identifies hand positions using an oriented hand bounding box. It also integrates a hand landmark model that processes the cropped hand bounding box provided by the palm detector and returns high-quality 2.5D landmarks.

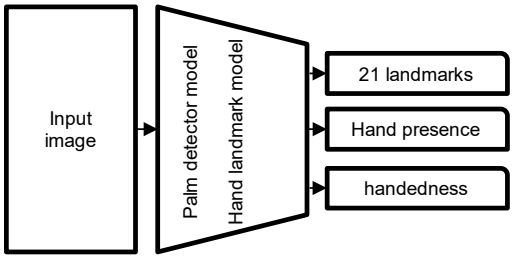



Fig. 3 Hand detection model

We use the output of the hand recognition model in Figure 3 to develop algorithms that track the user's interaction with our system. Since each task has its own unique way of being performed and requires specific instructions to be followed, the monitoring and evaluation algorithms are tailored to each specific task. Figure 4 shows the detailed process of how the algorithm works from start to finish.

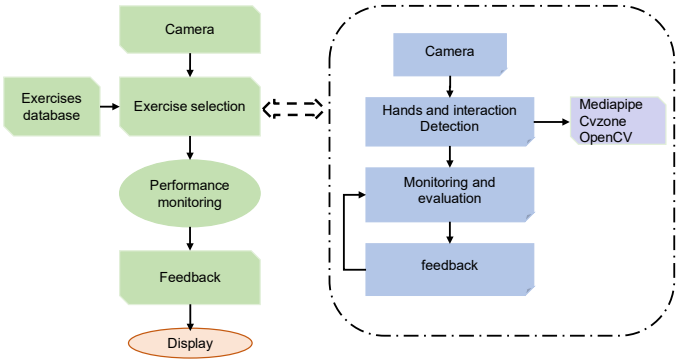



Fig. 4 Flow chart

When the system is started, the user must select a specific exercise from the database that the patient can perform. This step is necessary because our proposed system cannot evaluate the user's performance without first determining which task he will perform. After selecting a specific exercise from the database, the system extracts the rules of the task in preparation for evaluating the user's performance. To determine the quality of the exercises performed, we compare the recorded patient interaction with the therapist's instructions that were embedded in the system during the development of the prototype. Based on these instructions, the program provides recommendations on how to perform the exercise. We provide feedback that, if followed accurately, will enable the user to achieve the physical and cognitive goals defined by the therapist. The hand interaction coordinates obtained from the hand recognition model, as shown in Figure 2, are used to evaluate the patient's interaction.

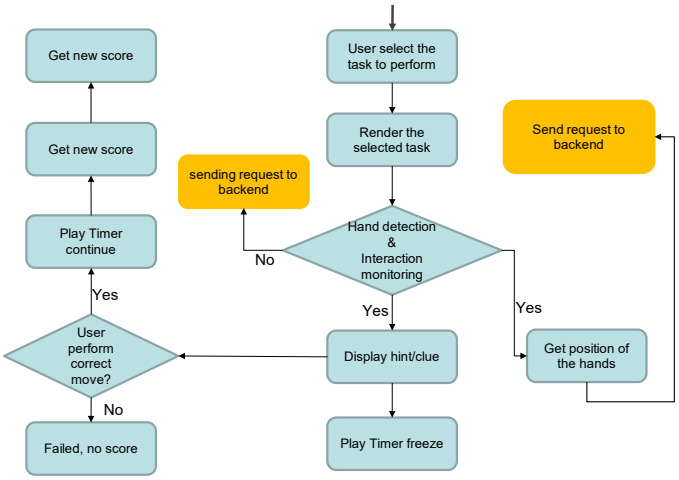



Fig. 5 System flow chart: Frontend

After a thorough examination of the system architecture and the algorithm of the proposed system, our attention will now turn to the frontend of the system. The frontend of the proposed system is the user-facing part of the system that allows patients to interact with it and perform rehabilitation exercises. It typically consists of a web application designed to be user-friendly and intuitive. The front-end application has several modules designed to support different aspects of cognitive rehabilitation, such as memory training, attention and concentration, problem solving, and decision making. These modules will include different exercises and activities tailored to the patient's specific needs and abilities. The front-end application will also include a user authentication and management system that allows patients to create and manage their accounts and track their progress. Figure 5 shows the step-by-step flowchart of the front-end system, while Figure 6 concludes with the overall system flowchart.

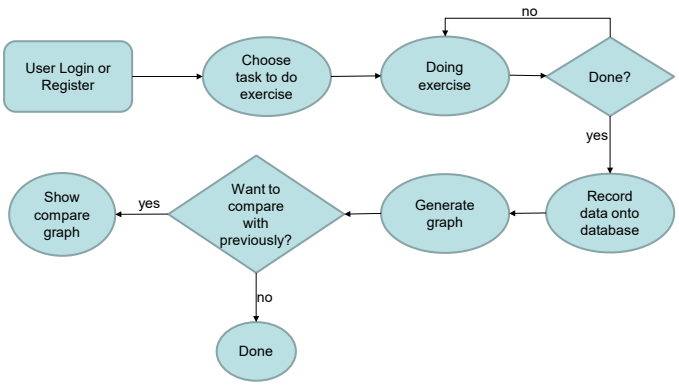



Fig. 6 System summary flow chart

## IV. TESTBED IMPLEMENTATION

The implementation of MEASURE involves several key components that work together to provide a comprehensive and effective exercise program. To demonstrate how the system works, let's take a closer look at the working prototype:

1. Central Control System: The central control system is responsible for managing the remote exercise equipment and monitoring each user's progress. This system includes software that can be accessed via a computer, allowing fitness professionals to remotely control the equipment and adjust the exercise program as needed.
2. User Interface: The user interface is the part of the system that users interact with directly. This interface is designed to be simple and intuitive, with clear instructions and feedback to guide users through each exercise.
3. Progress Tracking: The progress tracking feature of the system is designed to help users monitor their progress and stay motivated. The system tracks metrics such as heart rate, calories burned and distance traveled, and provides users with feedback and recommendations based on their performance.

Overall, the implementation of the multimodal exercise assistance system using remote equipment is a complex process that involves integrating multiple components into a cohesive and user-friendly system. Figure 9 below shows a user testing the working prototype to demonstrate how the system works.

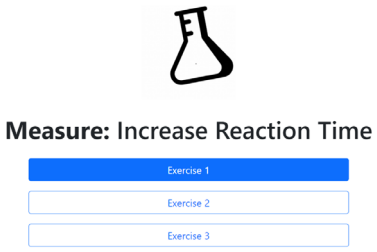



Fig. 7 Exercise selection


Fig. 8 Interaction with the selection Exercise

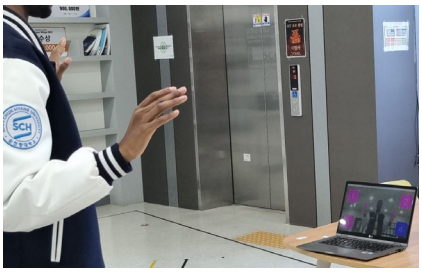


Fig. 9 User testing the prototype

## V. Discussion and Future Developments

Our system allows therapists to create individualized exercise schedules for patients based on their specific needs. This allows patients to receive cognitive rehabilitation from a convenient location without having to be physically present in the same room as the provider. The solution provides real-time exercise monitoring not only to the therapist, but also to every member of the rehabilitation chain, enabling them to track and maintain rehabilitation progress. In addition, patients can follow the therapist's exercise instructions and assignments by following the on-screen display. To further enhance the functionality of the system, we plan to implement additional technology, such as eye tracking, to facilitate navigation between menus. This is intended to increase user comfort by allowing them to operate the system while remaining in the same position in which they are performing their exercises. In addition, beyond simply following the movement of the fingers, we want to develop it into a form of measuring the movement of the joints of the upper body through human pose estimation for the upper body part.

## VI. Conclusions

The proposed system, MEASURE, has demonstrated its potential in assisting users to enhance their physical and cognitive abilities through virtual interactive games designed by therapists. Our system has shown reliable performance in detecting and providing feedback on correct posture during exercises, which can effectively encourage patients to reach their intended correct posture. The results of our study provide evidence that MEASURE can be a valuable tool in the rehabilitation process, as it provides personalized and interactive exercise routines that are tailored to each patient's needs. Further development and testing of the proposed system could lead to its widespread adoption and use in various rehabilitation and fitness applications. Overall, MEASURE is a promising solution that has the potential to revolutionize the way that people approach physical and cognitive rehabilitation, and we believe that it will have a positive impact on the lives of many individuals.

## Acknowledgment

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2023-2020-0-01832) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation). This research was also supported by Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (P0012724, The Competency Development Program for Industry Specialist). This work was also supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2019R1G1A1100699).

## References

[1] S. Vaughan, M. Wallis, D. Polit, M. Steele, D. Shum, and N. Morris, "The effects of multimodal exercise on cognitive and physical functioning and brain-derived neurotrophic factor in older women: a randomised controlled trial," *Age and ageing*, vol. 43, no. 5, pp 623-629, Sep. 2014.

[2] F. Lam, M. Huang, L. Liao, R. Chung, T. Kwok, and M. Pang, "Physical exercise improves strength, balance, mobility, and endurance in people with cognitive impairment and dementia: a systematic review," *Journal of physiotherapy*, vol. 64, no. 1 pp. 4-15, 2018.

[3] S. Vaughan, M. Wallis, D. Polit, M. Steele, D. Shum, and N. Morris. "Effects of physical exercise in older adults with reduced physical capacity: meta-analysis of resistance exercise and multimodal exercise," *International Journal of Rehabilitation Research*, vol. 40, no. 4 pp. 303-314, 2017.

[4] A. Daugherty, C. Zwilling, E. Paul, N. Sherepa, C. Allen, A. Kramer, C. Hillman, N. Cohen, and A. Barbey, "Multi-modal fitness and cognitive training to enhance fluid intelligence," *Intelligence*, vol. 66, pp. 32-43, 2018.

[5] T. Cordes, L. Bischoff, D. Schoene, and B. Wollesen, "A multicomponent exercise intervention to improve physical functioning, cognition and psychosocial well-being in elderly nursing home residents: a study protocol of a randomized controlled trial in the PROCARE (prevention and occupational health in long-term care) project," *BMC geriatrics*, vol. 19, pp. 1-11, 2019.

[6] M. Andriluka, S. Roth, B. Schiele, "Pictorial Structures Revisited: People Detection and Articulated Pose Estimation," *CVPR*, 2009.

[7] P. Fihl and T. B. Moeslund, "Pose estimation of interacting people using pictorialstructures," *7th IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 462–468, IEEE, 2010.

[8] W. Shih-En, V.Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," *IEEE conference on Computer Vision and Pattern Recognition*, pp. 4724–4732, 2016.

[9] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," *IEEE conference on computer vision and pattern recognition*, pp. 7291–7299, 2017.

[10] Y. Chen, Z. Wang, Y. Peng, Z. Zhang, G. Yu, and J. Sun, "Cascaded pyramid network for multi-person pose estimation," *IEEE conference on computer vision and pattern recognition*, pp. 7103–7112, 2018.

[11] K. Sun, B. Xiao, D. Liu, J. Wang, "Deep High-Resolution Representation Learning for Human Pose Estimation," *IEEE/CVF conference on computer vision and pattern recognition*, 2019

# Real-time Face Identification Using YOLO and CNN

Sejin Gown[1], Neunggyu Han[2*], Yunyoung Nam[3], and Seungmin Rho[4]
[1]*Department of Computer Science and Engineering, Soonchunhyang University, Asan 31538, Korea*
[2]*Department of ICT Convergence, Soonchunhyang University, Asan 31538, Korea*
[3]*Department of Computer Science and Engineering, Soonchunhyang University, Asan 31538, Korea*
[4] *Department of Industrial Security, Chung-Ang University, Seoul, South Korea*
*Contact: lovecein4858@naver.com

*Abstract* – **Face recognition technology is constantly evolving. Among them, 2D approaches reached some degree of maturity and reported very high rates of recognition. However, the 2D approach has a problem of poor performance depending on the ambient conditions and facial appearance. 3D approaches were proposed as an alternative solution to the problems mentioned above. The 2D approach requires no specific hardware, the 3D approach does. This paper conducted research on equipment with low performance. Therefore, a stable 2D approach that does not require specific hardware was judged to be suitable. First, a custom YOLO model is used to detect objects in the eyes, nose, and mouth. To extract feature values from detected objects, we use the VGG16 model with a reduced filter and input size. As a result of the test, it was possible to identify a specific person, but it was vulnerable to noise due to high-brightness images or changes in face angle. And there was a case of confusion with another person. To improve these problems, future research needs improvement.**

## I. Introduction

Adjabi et al. [13] wrote a review paper on face recognition. According to its contents, Face recognition technology is constantly evolving. Among them, 2D approaches reached some degree of maturity and reported very high rates of recognition. However, if the ambient conditions (e.g., lighting) or the facial appearance (e.g., pose or facial expression) change, this performance will degrade dramatically. 3D approaches were proposed as an alternative solution to the problems mentioned above. The advantage of 3D data lies in its invariance to pose and lighting conditions, which has enhanced recognition systems' efficiency. The 2D approach uses a camera. However, the 3D approach requires specific hardware devices. I conducted research on devices with low performance. In this paper, it was judged that a stable 2D recognition technology that does not require specific hardware is suitable. To implement face Identification, we need two deep learning models: feature extraction and object detection.

The feature extraction model mainly used VGG16 [1].
This paper is related as follows: Hridayami et al. [2] developed a real-time face recognition system by extracting human facial features. Dhuri et. al. [3] created a real-time parking occupancy detection system by extracting features of parking spaces.

YOLO [4] was used as a model for object detection. YOLO is a representative object detection model, and other models include R-CNN [5], Faster RCNN [6], and SSD [7]. Here, the R-CNN base uses a 2-stage detector, so it is not suitable for real-time detection. Rios et al. [11] compared yolo and ssd. SSD uses a 1-stage detector, but requires more epochs than YOLO and has lower accuracy. On the other hand, YOLO is a model suitable for real-time detection using a 1-stage detector. However, the VGG16 requires 15.5 GFLOPS, and the lightest of the YOLOV5 [8] requires 4.3 GFLOPS. Considering that the test environment Raspberry Pi is 13.5 GFLOPS, it is calculated at 0.7 fps, so the two models are unsuitable for use.

In this paper, we propose a method that can use face Identification while significantly reducing the amount of computation and capacity compared to existing face Identification models. First, to reduce the amount of computation of the VGG16 model, part of the face was trained instead of the whole face. It was assumed that features could be extracted normally even if the number of filters and image size were reduced. Therefore, a model with reduced input size and filter size was used. These models trained the left eye, right eye, mouth, nose, etc. Each of these models requires 20M FLOPS and a total of 80M FLOPS. A lightweight custom model based on the YOLOv5 model was used among the YOLO models. As the backbone, Darknet53-tiny used in YOLOv3-tiny [9] was used as the backbone.

In addition, CSPNET was applied to the C3 of the head. We will call this CSPC3. Finally, the size of the model was reduced to 300 MFLOPs by adjusting the number of filters. Therefore, in the Raspberry Pi environment, 36 fps was confirmed as a calculation result, and real-time performance is improved. In addition, the learning accuracy of VGG16 was more than 95%, and the Yolo model had 96% accuracy based on mAP 50.

## II. System Implementation

### A. Environmental Setup

TABLE.1 Experiment Environment

| | |
|---|---|
| Model | Raspberry Pi 4B |
| CPU | Cortex-A72 x4 |
| RAM | 4GB |
| OS | Ubuntu 22.04.1 LTS |

TABLE.2 Training Environment

| | |
|---|---|
| CPU | Xeon silver 4216 x2 |
| GPU | RTX A5000 x2 |
| RAM | 192GB |
| OS | Ubuntu 20.04.1 LTS |

### *B. Dataset*

Datasets were collected from Google search crawling and YouTube. Google search was used to collect the faces of various people. And YouTube data is less diverse, but it was used because it was possible to obtain data on various expressions and angles of the face. A total of 17324 datasets were used. First, manual labelling was performed on 200 datasets. After manual labelling, it was trained using YOLOv5x. Based on this model, 17324 data were automatically labelled. We checked the auto-labelling results and found nothing wrong. This labelling method has the advantage of being able to proceed faster than labelling all data manually. Finally, this data was used to train a lightweight custom model based on YOLOv5n.

Using a YOLOv5x-based learning model, 18073 left eyes, right eye, nose, and mouth datasets consisting of 10 classes were extracted from 10 YouTube videos. This data is used to train four VGG16 models used for feature extraction.

### *C. Model.*

We improved the speed by customizing the YOLOv5n model for object detection. YOLOv5n is a lightweight model. However, it was judged that it was still difficult to operate in the environment covered by this paper. First, the number of parameters was reduced by reducing the filter size. And Darknet53-tiny backbone using YOLOv3-tiny was used. In addition, the speed and performance were improved by applying the CSPC3 structure that applied CSPNet to the C3 layer. Fig.1 shows the structural diagram of the customized model.

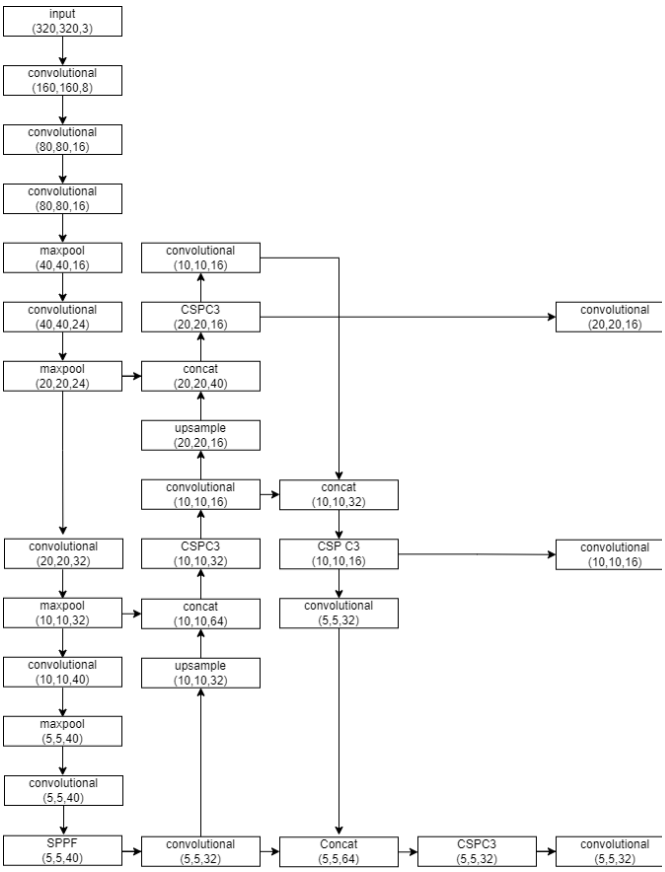


Fig.1 Custom Model Diagram

And the VGG16 was used for feature extraction. The existing VGG16FACE is a model for face recognition. The input image size is 224x224x3, and the number of filters is 64/128/256/512/512, which gradually increases. However, this paper deals with images of parts such as the eyes, nose, and mouth, which have fewer feature values than the face. Therefore, it is assumed that if a part of the face is used instead of the whole face, features can be extracted even if the number of filters and sizes are reduced. Therefore, the input image size was reduced to 96x96x1 and the number of filters was reduced to 4/8/16/32/32.

### *D. Implementation*

The trained VGG16 model was converted to TensorFlow Lite to optimize it for Raspberry Pi. TensorFlow Lite is a type of Tensorflow suitable for embedded environments. The output layer has been changed to enable feature extraction during the transformation process. We hanged it to flatten the layer, which is the middle layer, not the dense layer so that the value can be obtained from the middle layer. Using the existing output of the dense layer as is, we can only distinguish the 10 learned classes. However, by extracting feature values from the middle layer and comparing them, more classes can be distinguished. First, the data values of the results of inference through YOLO were

used. Confidence was set to a value of 65% or higher to reduce the number of misrecognized cases. Fig.2 shows the result screen.

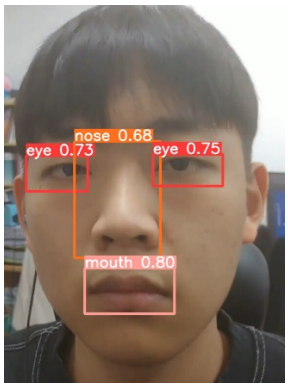



Fig.2 Detect Object

The image data extracted by the YOLO model is used to extract feature values using the VGG16 model. And all corresponding feature values are stored in storage as NumPy arrays. Two new feature values are judged using cosine similarity. Here, the cosine similarity is the degree of similarity between vectors measured using the cosine value of an angle between two vectors in the dot product space. Face Identification was implemented by finding and classifying classes with the highest similarity among those with cosine similarity of 0.85% or more.

## III. Experimental Results

### A. Model Comparison

Athreya et. al. [12] implemented face identification using yolov3 and VGG16 models. TABLE.3 compares the face identification and fps implemented in this paper and Athreya. Athreya's gpu is MX250 and Raspberry Pi 4B's GPU is Broadcom VideoCore VI. The performance of the two floating point can be seen in Table 5. Also, TABLE.4 is a table summarizing the performance of the YOLO model implemented in this paper.

TABLE.3 FPS for Each Model

| Model | Inference time (fps) |
|---|---|
| Athreya's Face Identification | 0.8 |
| This paper's Face Identification | 10 |

TABLE.4 Evaluation for Each Model

| Model | Accuracy (%) | | | Inference time (fps) |
|---|---|---|---|---|
| | eye | nose | mouth | |
| YOLOv5n | 98.8 | 98.8 | 99.2 | 5.26 |
| Customized Model | 94.9 | 96.5 | 96.7 | 18.1 |

TABLE.5 Floating Point Performance

| GPU | Performance (GFLOPS) | | |
|---|---|---|---|
| | FP16 | FP32 | FP64 |
| MX250 | 18.98 | 1,215 | 37.97 |
| Broadcom VideoCore VI | 64 | 32 | 8 |

### B. Experimentation

Three faces were registered with the camera attached to the Raspberry Pi. As a result of running the FACE Identification after registration, you can see the three faces are classified as Fig.4 Fig.5 Fig.6 images below.

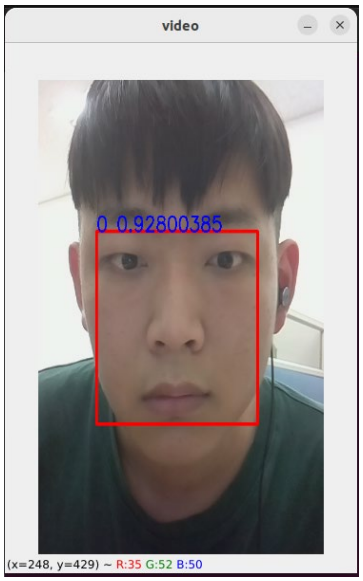



Fig.4 Face Identification No.0


Fig.5 Face Identification No.1

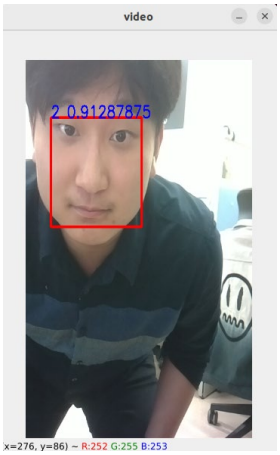


Fig.6 Face Identification No.2

## IV. Conclusions

This paper proposed a method that can perform face recognition even on low-performance devices. A face recognition model was trained using image data obtained from Google and YouTube. Face recognition was implemented by performing face detection and feature extraction with the two trained models. In the test environment of another paper, 0.8 fps came out, We have achieved a speed of 10 fps in my raspberry pi environment. And in an experiment to classify three people using a camera, 100% accuracy and a speed of 10 fps were achieved. However, problems were found that were vulnerable to noise and face angle change due to high brightness. In future studies, class expansion and utilization of more datasets are planned when learning the VGG16 model.

## Acknowledgment

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2023-2018-0-01799) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation) and the National Research Foundation of Korea(NRF) grand funded by the korea government(MSIT) (No. 2022H1D8A3038040)

## References

[1] Hridayami, P., Putra, I. K. G. D., & Wibawa, K. S. (2019). Fish species recognition using VGG16 deep convolutional neural network. *Journal of Computing Science and Engineering*, *13*(3). https://doi.org/10.5626/JCSE.2019.13.3.124

[2] Aung, H., Bobkov, A. v., & Tun, N. L. (2021). Face detection in real-time live video using yolo algorithm based on VGG16 convolutional neural network. *Proceedings - 2021 International Conference on Industrial Engineering, Applications and Manufacturing, ICIEAM 2021*. https://doi.org/10.1109/ICIEAM51226.2021.9446291

[3] Dhuri, V., Khan, A., Kamtekar, Y., Patel, D., & Jaiswal, I. (2021). Real-time parking lot occupancy detection system with VGG16 deep neural network using decentralized processing for public, private parking facilities. *Proceedings of the 2021 1st International Conference on Advances in Electrical, Computing, Communications and Sustainable Technologies, ICAECT 2021*. https://doi.org/10.1109/ICAECT49130.2021.9392506

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi.You only look once: Unified, real-time object detection. arXiv preprint arXiv:1506.02640,2015.

[5] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[J]. Computer Science,2013:580-587.

[6] Ren S,He K,Girshick R,etal.Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J].IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015:1-1.

[7] L. Wei, A. Dragomir: SSD: Single Shot MultiBox Detector.arXiv preprintarXiv:1512.02325v5,2016.

[8] Yao, J., Qi, J., Zhang, J., Shao, H., Yang, J., & Li, X. (2021). A real-time detection algorithm for kiwifruit defects based on yolov5. *Electronics (Switzerland)*, *10*(14). https://doi.org/10.3390/electronics10141711

[9] Xiao, D., Shan, F., Li, Z., Le, B. T., Liu, X., & Li, X. (2019). A target detection model based on improved tiny-yolov3 under the environment of mining truck. *IEEE Access*, *7*. https://doi.org/10.1109/ACCESS.2019.2928603

[10] Wang, C. Y., Mark Liao, H. Y., Wu, Y. H., Chen, P. Y., Hsieh, J. W., & Yeh, I. H. (2020). CSPNet: A new backbone that can enhance learning capability of CNN. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2020-June*. https://doi.org/10.1109/CVPRW50498.2020.00203

[11] Rios, A. C., dos Reis, D. H., da Silva, R. M., de Souza Leite Cuadros, M. A., & Tello Gamarra, D. F. (2021). Comparison of the YOLOv3 and SSD MobileNet v2 algorithms for identifying objects in images from an indoor robotics dataset. *2021 14th IEEE International Conference on Industry Applications, INDUSCON 2021 - Proceedings*. https://doi.org/10.1109/INDUSCON51756.2021.9529585

[12] Shet, Athreya V., et al. "Face Detection and Recognition in Near Infra-Red Image." 2022 6th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS). IEEE, 2022.

[13] Adjabi, Insaf, et al. "Past, present, and future of face recognition: A review." Electronics 9.8 (2020): 118

# An Equivalent Circuit of Microstrip Bend Discontinuity Element

Jeongho Park[1,*], Hyungzun Mun[2], Jongsik Lim[1,2], and Dal Ahn[1,2], Sang-Min Han[1], Seong-Ho Son[1], and Dong Min Kim[1]
[1]*Department of ICT Convergence, Soonchunhyang University*
[2]*Department of Electrical Communication and System Engineering, Soonchunhyang University*
*Contact: anotool0701@gmail.com, phone +82-10-2218-5922

*Abstract*—**This paper describes an equivalent circuit of microstrip bend element which is widely used in the design of microwave circuits. For the equivalent circuit, S-parameters obtained through EM (electromagnetic) simulation and simple circuit network analysis method are used instead of conventional methods using complex and difficult electromagnetic equations or numerical solutions. In the EM simulation process, the port feeding lines are de-embedded to produce the S-parameters of the bend discontinuity element itself. The S-parameters are converted into Z-parameters, which are applied for extracting the elements of the equivalent circuit. It is shown that the S-parameters of the proposed equivalent circuit are similar to those calculated from the electromagnetic simulation. Because the two S-parameters are similar to each other, the validity of the proposed equivalent circuit is verified.**

## I. Introduction

Microstrip transmission lines is widely used in microwave circuit design. Especially, bend discontinuity element is very frequently used when long microstrip lines are required to be curved or folded. There are parasitic components in microstrip bend elements, and they interferes signal transmission along microstrip lines [1]. Therefore, it is necessary to understand the characteristics of microstrip bend element, and to reflect them in circuit design.

Conventionally, equivalent circuits of microstrip bend were extracted by very difficult and complex electromagnetic equations or numerical analysis methods [2]. This study proposes a method to obtain the equivalent circuit, as the perspective of a circuit designer, by utilizing relatively simple network analysis method. The S-parameters obtained by electromagnetic simulation are converted into Z-parameters. If the equivalent circuit is analysed using basic circuit analysis theory, the equivalent circuit of bend elements can be obtained [3,4]. Therefore, it is relatively easy and simple to obtain the equivalent circuit of the microstrip bend elements without relying on conventional difficult methods.

## II. Microstrip Bend Discontinuity Elements

Fig. 1 shows a microstrip bend element, which is the most widely used in microstrip circuit design. P1 and P2 mean port1 and port2, respectively. $W_{50}$ is the width of the 50Ω microstrip line. The 50Ω microstrip transmission line with length $L_{de}$ is de-embedded during the extraction process of the equivalent circuit. This is because the equivalent circuit of only the pure discontinuity element part should be extracted. In Fig. 1, the input signal injected into P1 is transmitted to P2 through the bend discontinuity element.

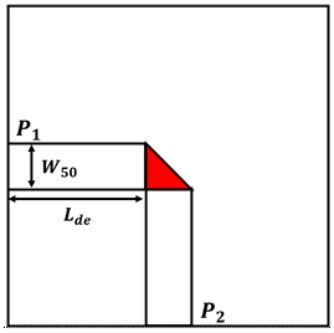



Fig. 1 Microstrip bend discontinuity element

Fig. 2 is the electric field distribution showing electromagnetic (EM) simulation results performed by Ansys HFSS (high frequency structure simulator). When microwave signal meets a bend discontinuity element, a distortion in the electromagnetic field distribution occurs. As the result, parasitic inductances and capacitances are generated additionally even they are small values. Normally, it is understood that the parasitic components do not exist in simple straight microstrip lines.

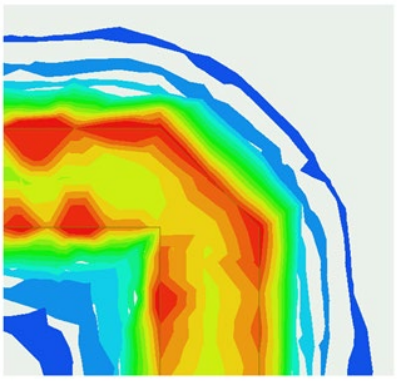


Fig. 2 Electric field distribution of microstrip bend discontinuity element

## III. Extracting Equivalent Circuit

Fig. 3 is the equivalent circuit of the microstrip bend discontinuity element proposed in this study. In the process of constructing the equivalent circuit, various topologies are possible according to physical interpretation. In this study, the equivalent circuit in Fig. 3 has been determined based on the signal transmission and fringing field described above. The extraction of the equivalent circuit is to find the element values of the circuit shown in Fig. 3.

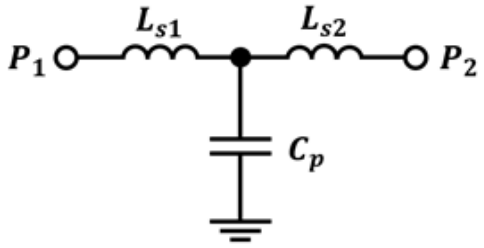



Fig. 3 Equivalent circuit of bend discontinuity element

It starts with the S-parameter ($[S]_{em}$) obtained from EM simulation to extract the equivalent circuit. From the S-parameters obtained from EM simulation, the port connecting lines are de-embedded to obtain the S-parameters of the discontinuity element only. Then, the S-parameters are converted into Z-parameters, and basic circuit analysis methods are applied to Fig. 3 to get the equivalent circuit elements. The equivalent circuit elements are calculated from Equations (1-a) ~ (1-c). Then, S-parameters ($[S]_{cir}$) can be obtained using circuit simulators to confirm the reliability of the proposed equivalent circuit. One can compare $[S]_{cir}$ with $[S]_{em}$. If a good similarity is observed between the two S-parameters, the proposed equivalent circuit of the bend discontinuity element is reliable.

$$L_{s1} = -\frac{im(Z_{11}) - im(Z_{12})}{2\pi f} \quad (1\text{-a})$$

$$L_{s2} = -\frac{im(Z_{22}) - im(Z_{12})}{2\pi f} \quad (1\text{-b})$$

$$C_p = \frac{1}{2\pi f \times im(Z_{12})} \quad (1\text{-c})$$

## IV. Verification of the Proposed Equivalent Circuit

In this study, a 50Ω microstrip line at the center frequency of 1.5 GHz was designed using the FR-4 substrate with the relative dielectric constant ($\varepsilon_r$) of 4.4 and the thickness of 0.8mm. The width of 50Ω line ($W_{50}$) is 1.48 mm. As shown in Fig. 1, microstrip feeding lines and discontinuity bend element were constructed and the corresponding EM simulation has been performed. Table 1 shows the extracted element values of the equivalent circuit for the microstrip bend.

Table 1. Equivalent circuit elements for the microstrip bend (@1.5GHz)

| Elements | Values |
|---|---|
| $L_{s1}$[nH] | 0.1564 |
| $L_{s2}$[nH] | 0.1564 |
| $C_p$[pF] | 0.2694 |

The 50Ω port connecting lines were connected to the equivalent circuit obtained by the proposed method, and the S-parameter ($[S]_{cir}$) was produced by Advanced Design System (ADS) from Keysight Technologies. One can compare $[S]_{cir}$ to $[S]_{em}$.

Fig. 4 shows the comparison between the two S-parameters. It can be seen that the S-parameters of the equivalent circuit well agree with the initially electromagnetically simulated S-parameters in terms of signal transmission and reflection coefficient.

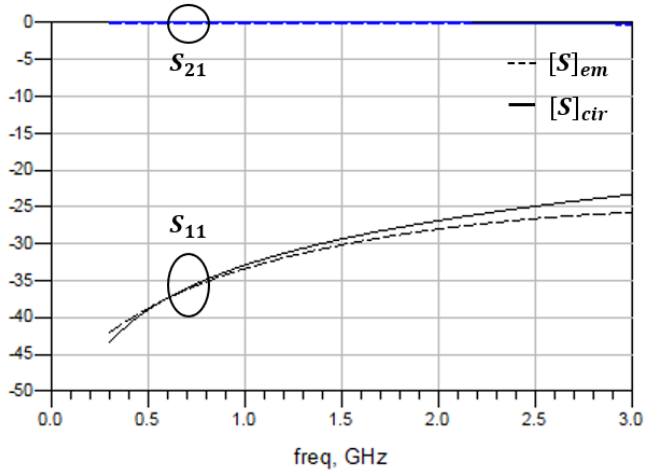



Fig. 4 Two S-parameters of the microstrip bend from EM and circuit simulations

## V. Conclusion

This work discussed the extraction of the equivalent circuit elements of microstrip bend discontinuity element which is widely used in microwave circuit design. Instead of previous difficult methods such as electromagnetic equations and numerical analysis, a relatively simple network analysis method using Z-parameters were adopted. The S-parameters of the extracted equivalent circuit showed a good agreement in those from the EM simulation. The proposed equivalent circuit extraction method is expected to be similarly applicable to other types of microstrip discontinuity elements.

## ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) (No. 2022H1D8A3038040), and ICAN(ICT Challenge and Advanced Network of HRD) program supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation) (IITP-2023-2020-0-01832) grant funded by the Korea government(Ministry of Science and ICT, MSIT).

## REFERENCES

[1] K. C. Gupta, *Microstrip Line and Slotlines* (2/3), Artech House Inc, MA, 1996, pp.196-200.

[2] W. Menzel, I. Wolff, "A Method for Calculating the Frequency-Dependent Properties of Microstrip Discontinuities", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 25, No. 2, pp. 107-112, Feb. 1977.

[3] D. G. Swanson and W. J. R. Hoefer, *Microwave Circuit Modelling Using Electromagnetic Field Simulation*, Artech House Inc, MA, 2003.

[4] D. M. Pozar, *Microwave Engineering* (4/e), John Wiley & Sons, NY, 2012, pp.174-178.

# Human Sleep Posture Classification using RGB and Thermal Imaging with Deep Learning Techniques

Awais Khan[1], Jung-Yeon Kim[2], and Yunyoung Nam[3]
[1]*ICT Convergence, Soonchunhyang University, Asan, South Korea*
[2]*ICT Convergence Research Centre, Soonchunhyang University, Asan, South Korea*
[3]*Department of Computer Science and Engineering, Soonchunhyang University, Asan, South Korea.*
*Contact: ynam@sch.ac.kr

*Abstract*— **Sleep monitoring, including accurate classification of sleep postures, is crucial for diagnosing and treating a range of sleep disorders, such as sleep apnea, restless leg syndrome, and rapid eye movement (REM) sleep behaviour disorder, which can have a significant impact on overall health and comfort. However, this task poses significant challenges due to the complex nature of the human body and variations that effect the human sleep. To address this challenge, we propose a novel approach for sleep posture classification by utilizing RGB and thermal cameras to capture color and heat information, respectively. This enables a comprehensive analysis of body position and potential discomfort during sleep. Our approach begins with capturing the dataset of sleep postures using an RGB and thermal camera in the form of videos with five different commonly used postures, including supine, left log, right log, prone left, and prone right, with 10 participants. In the first step, we normalize the database in a video frame, followed by selecting and modifying two pre-trained models, VGG16 and ResNet50, according to the dataset's nature. We trained both modified models using transfer learning and extracted the features, which are then passed to the Cubic SVM and Fine KNN for final classification.**

## I. INTRODUCTION

Sleep is essential for human body and brain, serving various biological functions [1][1]. Sleep research is broadly classified into two categories. One category monitors physiological variables such as ECG and EEG during sleep, which are employed to classify sleep stages and evaluate the quality of sleep [2], [3], and [4] [2-4]. The other category observes the external body behaviour during sleep, including posture and movement detection [5, 6]. One may determine the quantity and quality of a sleep pattern by monitoring sleep position and movement. Human sleep includes periods of immobility as well as movements such position shifts or sleep poses. According to research, the amount of sleep and movements of the body are directly related [7]. Major movements, where the entire body changes position, suggest lighter sleep and the body may be preparing to wake up. Sleep apnea and snoring, restless leg syndrome (RLS), rapid eye movement (REM) and periodic limb movement (PLM) sleep behaviour disorder (RBD) are three sleep disorders connected to body movement and position changes. Polysomnography (PSG), which measures EEG, electromyography (EMG), electrooculography (EOG), breathing-related variables, and movement signals, is the gold standard approach for evaluating sleep. Despite PSG's usefulness in providing valuable information, it is difficult, time-consuming and expensive. Furthermore, patients must sleep in a laboratory setting, which may compromise sleep quality, and the requirements and characteristics of PSG make it unsuitable for the long-term monitoring. The actigraphy is an alternative diagnosis method that requires wearing a device. However, some actigraphy users may experience anxiety or panic, especially children [8].

Problems of conventional approaches may be overcome by quantitative non-contact sleep monitoring techniques, which may also be used to identify sleep disorders, improve sleep, and improve quality of life. Video-based techniques, sensor fusion and instrumented mattresses are the three groups of non-contact techniques for posture recognition. In other techniques, instrumented pillows and mattresses are utilized to study sleep positions [9, 10]. Lee et al. [11] identified different sleep postures using Kinect v2 skeleton tracking, which measures the x, y, and z spots of 25 joints. However, some non-contact sleep monitoring approaches require patients to avoid any bed covering (blankets or sheets), which may be inconvenient [12, 13]. Other approaches make use of fusion sensors, which combine a depth camera and an instrumented pressure mattress, to evaluate data and detect sleeping positions automatically. However, these methods are complex and expensive, limiting their practicality [14, 15].

The objective of this study is to propose a transfer learning approach for feature extraction from pre-trained CNN models (e.g., VGG16 and ResNet50) and modify the fully-connected layer of the classifier block. The proposed method is applied to the human sleep posture RGB and thermal camera datasets to train the modified classifier block. Finally, the extracted features are fed to cubic SVM and fine KNN for the final classification of human sleep postures.

## II. METHODOLOGY

Our novel approach to non-contact sleep posture monitoring is based on capturing the dataset of sleep postures using an RGB and thermal camera, in the form of videos of ten participants in five commonly used postures, including supine, left log, right log, prone left, and prone right. In order to ensure robustness and prevent overfitting, we first normalized the database in a video frame and then balanced the data. Next, we selected and modified two pre-trained models, VGG16 and ResNet50, these CNN models that have been widely used in various computer vision tasks. These models are known for their high accuracy and strong feature extraction capabilities. In addition, they have large numbers of layers and parameters, which enable them to learn complex features from large datasets in less time. Using transfer learning, we trained both modified models and extracted their features, which were then passed to the cubic SVM and fine KNN for final classification. This comprehensive approach allows for accurate and reliable classification of sleep postures without requiring any contact with the patient, potentially revolutionizing the diagnosis and treatment of sleep disorders.

### *A. Participant Recruitment*

Ten healthy male adults were recruited from both the university and community for this study, with an average age of 24 years. The participants had an average height of 177 cm and weight of 60.6 kg. To ensure that the deep learning model was robust and generalizable, participants of varying weights and heights were recruited. The study aimed to gather a diverse range of participants in order to adequately test the efficacy of the deep learning model. It is important to note that participant recruitment was conducted ethically and with informed consent. The study adhered to all relevant guidelines and regulations.

### *B. Data Collection*

In this study, the thermal and RGB camera was positioned 1.9 meters above the bed. The RGB camera records the image with the resolution of 464 X 848 pixels and with 30 frame per second and as well as the thermal camera records the image with the resolution of 624 X 832 pixels and with 8.2 frame per second. Ten participants were instructed to assume five different sleep postures, namely supine, left log, right log, prone left, and prone right. Participants were given time to adjust to their most comfortable position before data collection. The participants maintained their posture during data collection for ten to fifteen second, which was collected continuously throughout the experiment. The posture of each participant was evaluated by a selection of 300 images. The sample of dataset is shown in Figure1.

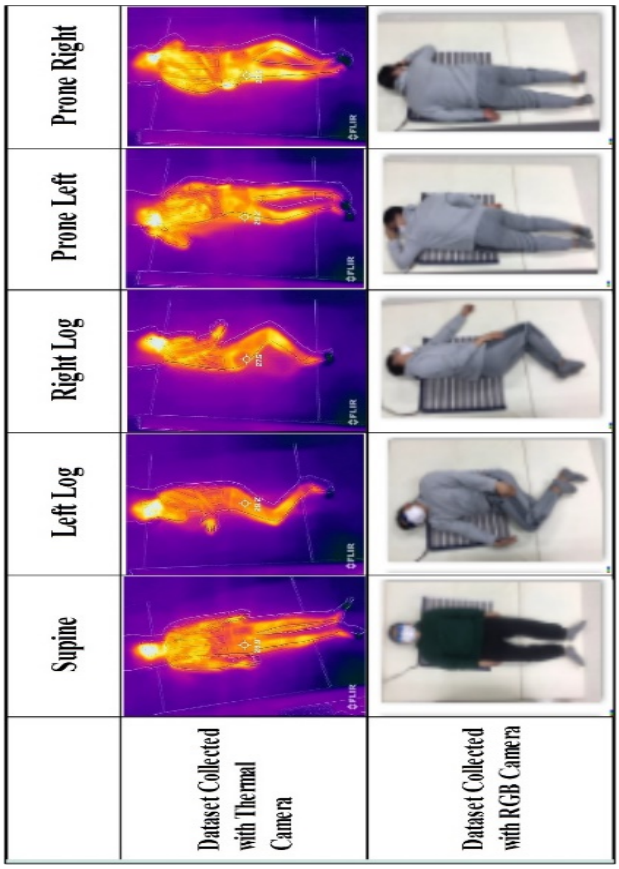



Figure 1. Sample sleep posture dataset of thermal as well as RGB images.

### *C. Modified VGG16 and Resnet50*

Convolutional Neural Networks (CNNs) are a successful deep learning technique used in image recognition, classification, and object detection. VGG16 is a CNN model that focuses on using identical pooling and convolutional layers with fewer hyperparameters. It comprises 13 convolutional layers and three fully connected layers, totalling 16 layers, and was initially trained on the ImageNet dataset. In this work, the VGG16 model is modified through transfer learning by removing the last fully connected layer and adding a new one that classifies images supine, left log, right log, prone left, and prone right classes. Transfer learning is employed to train the modified model on a selected sleep posture dataset. Features are extracted from the seventh fully connected layer, and a vector of dimension is obtained. The network architecture is visually represented in Figure 2.

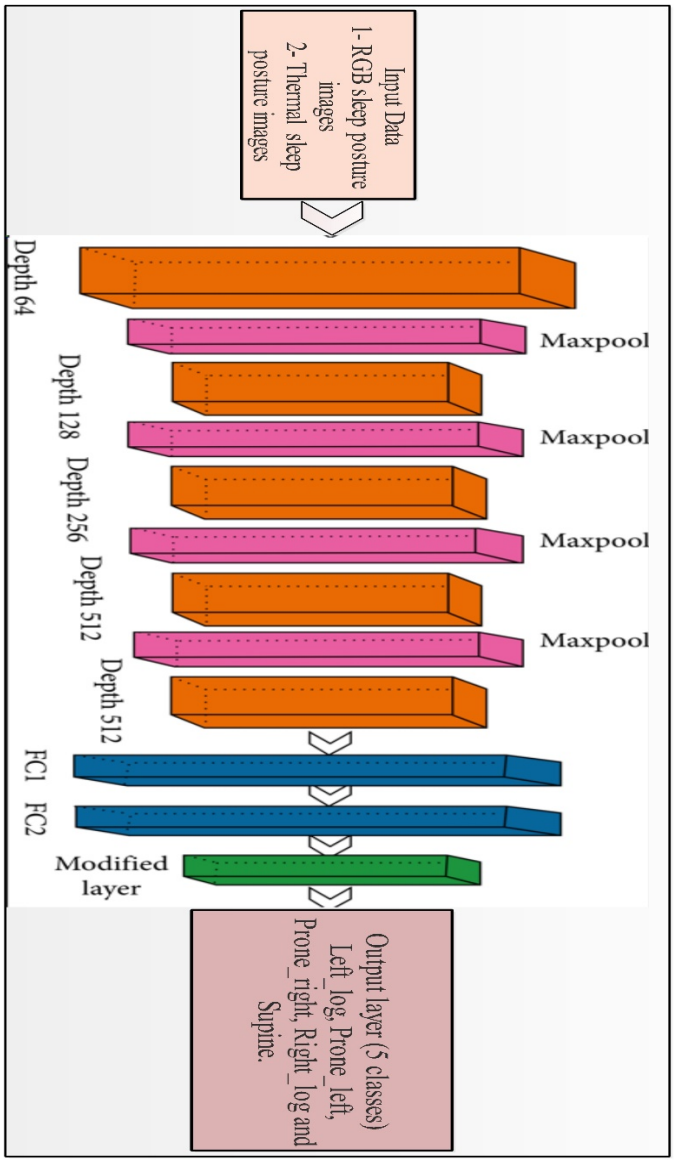



Figure 2. Modified architecture of VGG16.

### *D. Modified Resnet50*

The ResNet50 model consists of 50 layers, including convolutional and fully connected layers, and was originally trained on the ImageNet dataset. The model is modified using transfer learning by removing the last fully connected layer and replacing it with a new one that classifies images into supine, left log, right log, prone left, and prone right classes. The modified model is trained on a selected sleep posture dataset using transfer learning. During training, features are extracted from the fully connected layer, resulting in a vector of dimension 2048.

## III. RESULTS AND DISCUSSION

This section presents the results of an experimental study conducted on the sleep posture dataset. The dataset was divided into training and testing sets using an 80:20 split ratio. The training images are 14457 and the test images utilizes are 2891. The training process utilized a Stochastic Gradient Descent (SGD) optimization algorithm with 5 epochs, a mini-batch size of 8, and a learning rate of 0.0001. To evaluate the performance of the classifiers, various metrics including recall rate, precision, and accuracy were utilized in a 5-fold cross-validation. All simulations were conducted in MATLAB 2022a on a computer with a Core i7 processor and 8 GB of RAM.

### *A. Experimental results on RGB image dataset*

The results of sleep posture dataset with RGB camera using the proposed method are shown in Table I. Cubic SVM (CSVM) and fine KNN (FKNN) are utilizes for the classification results. The best accuracy is achieved on the CSVM of 94.2% while using the VGG16 model with the computational time of 49.5 sec, whereas the precision rate is 94.18 and the recall rate was 94.3. However, the FKNN got the accuracy of 92.8%, whereas the precision rate is 92.8 and the recall rate was also 92.8 and with the computational time of 69.8 sec which is higher than the CSVM. By using the resnet50 model the CSVM achieved the accuracy of 93.7% but the computational time was less as compared to the VGG16 model that is 24.3 sec, with the precision rate of 93.7 and the recall rate was 93.7.

Table II. Proposed results VGG16 and ResNet50 model using the RGB image dataset.

| DL Models | ML Classifiers | Accuracy | Time (sec) | Precision Rate | Recall Rate |
|---|---|---|---|---|---|
| **VGG16** | 1- CSVM | 94.2% | 49.5 | 94.18 | 94.3 |
| | 2- FKNN | 92.8% | 69.8 | 92.8 | 92.8 |
| **ResNet-50** | 1- CSVM | 93.7% | 24.3 | 93.7 | 93.7 |
| | 2- FKNN | 93.3% | 34.16 | 93.3 | 93.4 |

### *B. Experimental results on thermal image dataset*

The results of sleep posture dataset with RGB camera using the proposed method are shown in Table III. The classification task was performed using both cubic SVM and fine KNN algorithms. The highest accuracy of 96.2% was achieved using the VGG16 model, with a computational time of 55 seconds. The corresponding precision and recall rates were 96.6 and 96.5, respectively. On the other hand, the fine KNN algorithm yielded an accuracy of 92.8%, with a precision and recall rate of 96.3, and a computational time of 70.2 seconds, which was higher than that of the cubic SVM. By utilizing the ResNet50 model, the cubic SVM achieved an accuracy of 95.2%, with a computational time

of 24.7 seconds, which is less than that of the VGG16 model. The corresponding precision and recall rates were both 95.9.

Table IV. Proposed results VGG16 and ResNet50 model using the thermal image dataset.

| DL Models | ML Classifiers | Accuracy | Time (sec) | Precision Rate | Recall Rate |
|---|---|---|---|---|---|
| VGG16 | 1- CSVM | 96.2% | 55 | 96.6 | 96.5 |
| | 2- FKNN | 96% | 70.2 | 96.3 | 96.3 |
| ResNet-50 | 1- CSVM | 95.2% | 24.7 | 95.9 | 95.9 |
| | 2- FKNN | 94.8% | 32.9 | 95.2 | 95.2 |

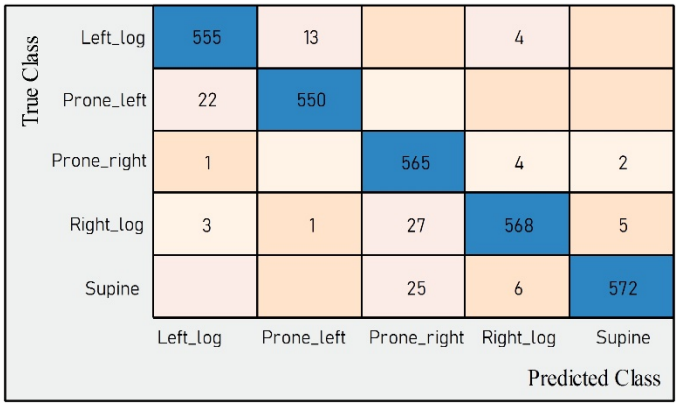



Figure 3. Proposed confusion matrix for VGG16 model using the thermal image dataset.

## IV. Conclusions

In this research paper, we propose a CNN and deep learning approach to accurately detect various sleep postures using RGB and thermal images. To this end, we conducted experiments using our proposed method on a simulated sleep dataset. The results of the experiments demonstrate that our method achieves better accuracy for thermal images compared to RGB images, with an average accuracy of 96.2% and 94.2%, respectively, after pre-processing the data. These findings suggest that thermal images can be more effective than RGB images in detecting body positions during sleep. Our approach can help develop low-cost and non-intrusive real-time sleep monitoring systems and benefiting healthcare professionals.

As future research directions, we plan to explore an intelligent non-contact monitoring system that utilizes the same setup for both sleep labs and home environments. Investigating the effect of using the CNN by adding more classes with different blanket conditions is another ongoing topic of interest. Furthermore, we aim to compare our approach with other methods such as actigraphy to better understand the relationship between sleep postures, movements, and sleep quality.

## Acknowledgment

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2023-2020-0-01832) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation).

## REFERRENCES

[1] S. S. Choudhary and S. R. Choudhary, "Sleep effects on breathing and respiratory diseases," *Lung India: official organ of Indian Chest Society,* vol. 26, no. 4, p. 117, 2009.

[2] S. M. Mohammadi, S. Kouchaki, M. Ghavami, and S. J. J. o. n. m. Sanei, "Improving time–frequency domain sleep EEG classification via singular spectrum analysis," *Journal of neuroscience methods,,* vol. 273, pp. 96-106, 2016.

[3] S. Kouchaki, S. Sanei, E. L. Arbon, D.-J. J. I. T. o. N. S. Dijk, and R. Engineering, "Tensor based singular spectrum analysis for automatic scoring of sleep EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,,* vol. 23, no. 1, pp. 1-9, 2014.

[4] S. M. Mohammadi, S. Enshaeifar, M. Ghavami, and S. Sanei, "Classification of awake, REM, and NREM from EEG via singular spectrum analysis," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 4769-4772: IEEE.

[5] C.-H. Kuo, F.-C. Yang, M.-Y. Tsai, M.-Y. J. B. E. A. Lee, Basis, and Communications, "Artificial neural networks based sleep motion recognition using night vision cameras," *Biomedical Engineering: Applications, Basis and Communications,,* vol. 16, no. 02, pp. 79-86, 2004.

[6] M. Blagrove, D. Owens, I. MacDONALD, N. Sytnik, P. Tucker, and S. J. J. o. s. r. Folkard, "Time of day effects in, and the relationship between, sleep quality and movement," *Journal of sleep research* vol. 7, no. 4, pp. 233-239, 1998.

[7] J. Van Hilten *et al.*, "Nocturnal activity and immobility across aging (50–98 years) in healthy persons," *Journal of the American Geriatrics Society,* vol. 41, no. 8, pp. 837-841, 1993.

[8] S. M. Mohammadi, M. Alnowami, S. Khan, D.-J. Dijk, A. Hilton, and K. Wells, "Sleep posture classification using a convolutional neural network," in *2018 40th Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, 2018, pp. 1-4: IEEE.

[9] E. Hoque, R. F. Dickerson, and J. A. Stankovic, "Monitoring body positions and movements during sleep using wisps," in *Wireless Health 2010*, 2010, pp. 44-53.

[10] T. Harada, A. Sakata, T. Mori, and T. Sato, "Sensor pillow system: monitoring respiration and body movement in sleep," in *Proceedings. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000)(Cat. No. 00CH37113)*, 2000, vol. 1, pp. 351-356: IEEE.

[11] M.-C. Yu, H. Wu, J.-L. Liou, M.-S. Lee, and Y.-P. Hung, "Multiparameter sleep monitoring using a depth camera," in *Biomedical Engineering Systems and Technologies: 5th International Joint Conference, BIOSTEC 2012, Vilamoura, Portugal, February 1-4, 2012, Revised Selected Papers 5*, 2013, pp. 311-325: Springer.

[12] X. Liu and S. Payandeh, "Toward study of features associated with natural sleep posture using a depth sensor," in *2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2016, pp. 1-6: IEEE.

[13] V. Metsis, D. Kosmopoulos, V. Athitsos, F. J. P. Makedon, and u. computing, "Non-invasive analysis of sleep patterns via multimodal sensor input," *Personal and ubiquitous computing,* vol. 18, pp. 19-26, 2014.

[14] M. Papakostas, J. Staud, F. Makedon, and V. Metsis, "Monitoring breathing activity and sleep patterns using multimodal non-invasive technologies," in *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, 2015, pp. 1-4.

[15] G. E. Hinton and R. R. J. s. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science,* vol. 313, no. 5786, pp. 504-507, 2006.

# Finding Face in Children's Drawings using U-Net Segmentation Techniques

*Vesal Khean[1], Chomyong Kim[2], Yunyoung Nam[3]*

*[1]Department of ICT Convergence, Soonchunhyang University, Asan, 31538, Republic of Korea*

*[2] ICT Convergence Research Centre, Soonchunhyang University, Asan, South Korea*

*[3] Department of Computer Science and Engineering, Soonchunhyang University, Asan, South Korea*

**Contact: ynam@sch.ac.kr*

*Abstract*— **Analyzing children's drawings can provide valuable insights into their emotional development and well-being. Children between the ages of 3 and 8 always enjoy drawing, which can express their feelings and development. Especially, the study focuses on social situations in which teachers respond to negative emotional expressions and socially charged acts of children, which are characterized by feelings of anger, irritation, and distress. However, manually analyzing these drawings is time-consuming and subject to human biases. This paper introduces a new deep-learning approach for detecting faces in children's drawings. The first step collects a dataset of children's drawings. We will use the U-Net model for image segmentation and OpenCV to find faces from children's drawings. Our proposed approach can provide an efficient and accurate to automatically identify expression faces in children's drawings. Additionally, this study contributes to the practical applications in various fields, including education, and children's welfare.**

## I. INTRODUCTION

Children's drawings have been recognized as a valuable tool for understanding their emotional interaction. However, analyzing these drawings manually can be time-consuming and subjective [1]. In recent years, computer vision techniques have advanced appreciably, and one area where techniques have displayed great potential is in the analysis of children's drawings. In particular, the ability to automatically detect and segment faces in children's drawings can have a wide range of applications, from educational research to artistic analysis [6].

The development of deep learning has enabled the development of image segmentation models that can accurately detect and find different objects and patterns within images. In this context, the use of U-Net segmentation techniques has demonstrated effectiveness in solving the problem, that utilizes U-Net segmentation deep learning to automatically detect and find the face of the express feeling, emotional distress in children's drawings [5]. The U-Net convolutional neural network architecture was initially created for the purpose of segmenting biomedical images. Its unique architecture includes an encoder and a decoder that allow for high-resolution segmentation results [8], even when working with limited training data. By adopting this architecture to the task of detecting faces in children's drawings, researchers have been able to achieve good results [7].

U-Net segmentation techniques offer a significant advantage in analyzing large datasets of children's drawings within a short period, making it a key benefit in this proposed. Additionally, this technique can be applied to drawings produced by children with developmental disabilities.

In this study, has to provide a fast and objective tool for analyzing the children's drawings, importantly reducing the time for analyzing the emotion indicator in children, we proposed the idea of using deep learning algorithms to automatically detect and find different faces in children's drawings. Moreover, which can be useful in various applications, including education, healthcare, and related fields.

## II. RELATED WORK

Various approaches for analyzing for human figure drawings as a diagnostic tool for emotional distress in children has been extensively studied [7]. For example, a study by Lawrie, Nicholas R (2022), the study discovered that interest of researchers in drawings stems from their ability to uncover emotional distress in children who may lack the vocabulary or self-awareness to sufficiently convey their emotional state. Despite having been a subject of study in computer science for many years, machine learning is not extensively utilized in the social sciences because its primary focus is on accurately predicting out-of-sample data rather than explaining the causal relationships between endogenous and exogenous variables. Other studies have focused on analyzing on the object mouth, lips shape, and lips colors [3], for example, a study by Ju, Zhen, Xiang Lin, Fangqi Li, and Shilin Wang (2018) found that the visual information of the lips can improve the effectiveness of speech recognition and identity verification that is based on lip movements. Lip information in visual data encompasses factors such as lip shape, segmentation, and color. Precise segmentation of the lips is crucial for applications that rely on this information. The last one of our previous studies have explored about Concentration is crucial for children to nurture good personality and develop well. To observe teachers and children in concentration-oriented preschool education activities [6], for example, a study by Jin, Ting, Zhuang Ma, Jinfang Niu, and Peng Su (2022) found that the image segmentation and target of extraction of preschool

education activity space for improving children's concentration.

## III. METHODOLOGY

### A. Image Processing

The proposed approach for finding face from children's drawing involves several steps of image processing techniques. First, the input drawing image is preprocessed to enhance its quality and remove noise such as resizing, smoothing, and histogram equalization. Then, the preprocessed image is segmented using U-Net deep learning architecture to find the different objects and patterns in the image. Finally, the finding face of the drawing is determined based on the predicted and find the face of mask assigned to the segmented objects.

We used OpenCV library for various image processing tasks, the first step was to read the image files and pre-process them to remove any noise, resizing, and converting images to grayscale.

### B. Create Mask

Creating a mask has a lot of tools for creating mask images. In this study, we used Procreate app to work on this task. Creating a mask image in Procreate is an essential step in preparing our image for the segmentation model. A mask image is a black and white image that defines which parts of our image should be included or excluded from the segmentation model.

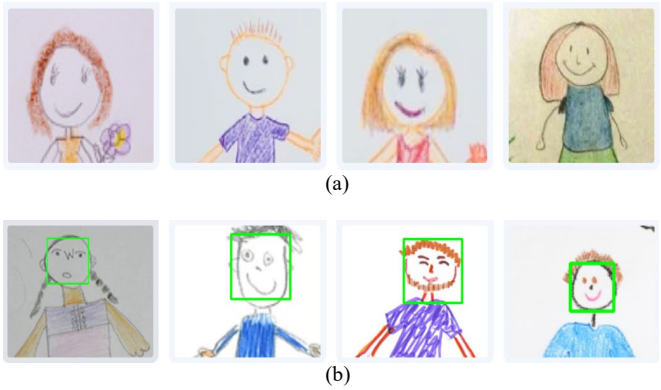
### C. U-Net Model

We implement the U-Net model for image segmentation, which is a popular deep learning architecture that has been widely used for image segmentation tasks. Comprised of an encoder and decoder network, the U-Net architecture is linked by a bottleneck layer that condenses feature maps. The encoder network performs down sampling of the feature maps, while the decoder network performs up sampling to generate the segmentation mask. The U-Net architecture has been shown to achieve high accuracy and efficiency in image segmentation tasks.

To train the U-Net model, we use a dataset of children's drawing. The dataset was partitioned into a training set and a test set, where the binary cross-entropy loss function and Adam optimizer were utilized. The model is trained for 100 epochs and achieves a validation accuracy of 96%.

One of the challenges faced during this proposed was the use of the dlib python library that used in computer vison to detect the face in children's drawing. The results were not good in that call did not detect all the faces all the children's drawings images, due to the variations in the drawing and the lack of accurate features for detection in Fig. 1. As a result, an alternative approach was used to find the face area first by segmenting the images using the U-Net model. This approach was found to be more accurate and reliable in detecting the face area in children's drawing.

The U-Net segmentation model proved to be effective in precisely detecting the face area in children's drawing, which is an essential step in analyzing the face depicted in the drawing.



(a)

(b)

Fig. 1 Sample display detecting face about using Dlib library.

In Fig.1 The first image (a) and the second image (b), are collected from children's drawings. So, the result shows that in the image (a) divide the image with non-detected and in image (b) provide the image with detected by using Dlib.

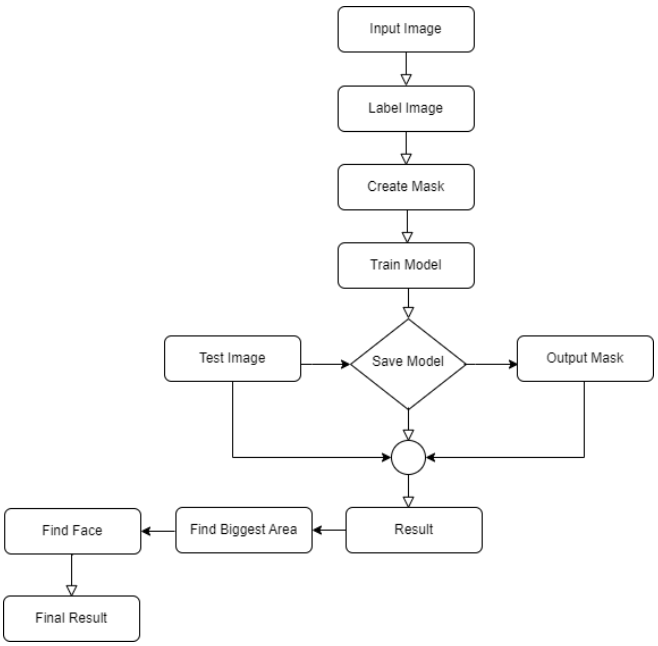



Fig. 2 Flowchart of the proposed method for finding face.

## IV. EXPERIMENTAL RESULTS

To evaluate the performance of our proposed approach for detect face from children's drawing using U-Net segmentation deep learning, we conducted an experiment using a dataset of 90 images collected from Google Images. We randomly divided the dataset into a train set of 61 images and a test set of 29 images.

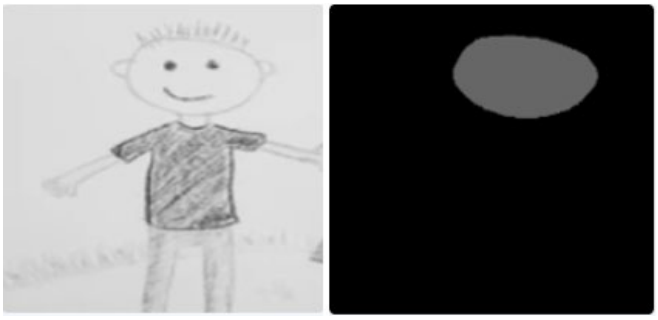
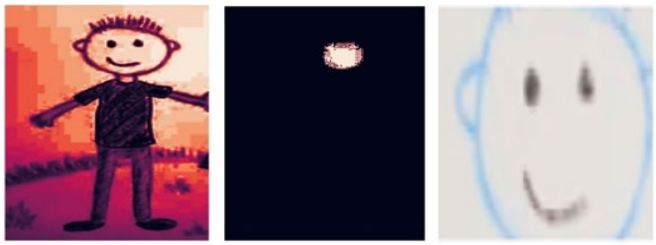


Fig. 3 The sample of the creating mask image using procreate website.

In Fig.4 In the input image, we have removed the background and find the biggest area of the face.

(a) (b) (c)

Fig. 4 The sample of the (a) input image along with its corresponding (b) ground truth mask and (c) predicted mask side by side for visual comparison and evaluation of the model's performance.



(a) (b) (c)

Fig. 5 The sample of (a) input image along with its corresponding (b) biggest area and (c) predict finding face result.

The experiment results are evaluated on the children's drawing dataset. The U-Net model training process is used for finding a face in children's drawings performance. The detection results are shown in Fig 6.


Fig. 6 Visualize the training and validation accuracy and loss during the model training process.

## V. Conclusions

In conclusion, experiments have demonstrated that implementing U-Net can be an effective method for facial segmentation. we proposed a novel approach for finding faces from children's drawings using U-Net segmentation deep learning. The proposal involves training a U-Net model on a dataset of images and then using this model to automatically detect and find the different faces in children's drawings. Our approach is to provide a fast and objective tool for analyzing the face of children, which can be useful in education, and healthcare.

We presented the details of our proposed approach, including the dataset used, the training process, and the experimental results. Our experiment results demonstrated that our proposed approach achieved an overall accuracy of 96%. These results suggest that our approach using U-Net segmentation deep learning can effectively find the face in children's drawings. Nevertheless, future research needs to address certain limitations in our approach, such as the requirement for larger and more diverse datasets to enhance the model's generalizability. Finally, we believe that our approach represents an importance step towards using deep learning for analyzing children's faces and making a positive impact in various domains.

In addition to the U-Net segmentation model, we are planning to conduct research on the facial area and eyes, and lips to make emotion in children.

## Acknowledgment

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) program(IITP-2023-2020-0-01832) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation)

## References

[1] Moula, Zoe, Nicola Walshe, and Elsa Lee. "Making nature explicit in children's drawings of wellbeing and happy spaces." *Child Indicators Research* 14 (2021): 1653-1675.

[2] Jolley, Richard P., and Grégory Dessart. "Emotional Expression in Children's Drawings of God." In *When Children Draw Gods: A Multicultural and Interdisciplinary Approach to Children's Representations of Supernatural Agents*, pp. 247-284. Cham: Springer International Publishing, 2023.

[3] Ju, Zhen, Xiang Lin, Fangqi Li, and Shilin Wang. "Lip segmentation with muti-scale features based on fully convolution network." In *2018 IEEE Third International Conference on Data Science in Cyberspace (DSC)*, pp. 365-370. IEEE, 2018.

[4] Lu, Yuanyao, and Qingqing Liu. "Lip segmentation using automatic selected initial contours based on localized active contour

[5] Jin, Ting, Zhuang Ma, Jinfang Niu, and Peng Su. "Image Segmentation and Target Extraction of Preschool Educational Activity Space for Improving Children's Concentration." *Traitement du Signal* 39, no. 5 (2022).

[6] Cekaite, Asta, and Anna Ekström. "Emotion socialization in teacher-child interaction: Teachers' responses to children's negative emotions." *Frontiers in Psychology* 10 (2019): 1546.

[7] Lawrie, Nicholas R. "Using Machine Learning to Analyze Children's Drawings as Indicators of Mental Well-Being." (2022).

[8] Beltzung, Benjamin, Marie Pelé, Julien P. Renoult, and Cédric Sueur. "Deep learning for studying drawing behavior: A review." *Frontiers in Psychology* 14 (2023).

# A Study on Analyzing the Sentiment of Monologues by Children through Supervised and Unsupervised Learning Techniques

Ahsan Aziz[1], Chomyong Kim[2] ,and Yunyoung Nam[3]
[1]Department of ICT Convergence, Soonchunhyang University, Asan, 31538, Republic of Korea
[2]ICT Convergence Research Centre, Soonchunhyang University, Asan, South Korea
[3]Department of Computer Science and Engineering, Soonchunhyang University, Asan, South Korea
*Contact: ynam@sch.ac.kr

***Abstract*— In this paper, we propose an innovative approach which is to analyze the sentiment in the text using the basic emotions as predefined four classes which are (Happy, sad, surprised, and angry). In the following paper, we also aim to classify responses that we get from YouTube. The goal of this study is to provide a more exact and accurate understanding of the children's emotions and then to provide more effective support for the children who have experienced abuse, targeted interventions should be implemented to address their specific needs and promote their sense of comfort. To accomplish this task, we employ natural language processing techniques, including the bag-of-the-words approach, to extract emotions from the text data. In the first step, we have to assemble the data in the arrangement of videos and audio files and convert the speech into text then we clean the data and pre-process the text data by removing stop words, punctuations, and other irrelevant text, and then tokenize the text into individual words after that the process of feature extraction was implemented and then in the final section applied the unsupervised approach module called VADER(valence Aware Dictionary and Sentiment Reasoner) which is lexicon-based approach and also classifying four classes (Happy, sad, surprised, and angry). Results indicate that our approach identifies the emotions expressed by children who get abused and provides valuable information on the prevalence of different emotional states to the teacher and the parents. This study contributes to the growing literature on natural language processing and machine learning in the field of psychology and the emotional feelings of children.**

## I. INTRODUCTION

Emotions play a crucial role in the emotional well-being of children, influencing their cognitive, social, and behavioural development. Emotions can be expressed through various channels, including verbal and non-verbal cues such as text. Children, especially those in school settings, may experience a range of emotions, including happiness, sadness, surprise, and anger, in response to various situations.

One significant aspect to deliberate in the situation of children's emotions is the occurrence of abuse in schools. School abuse refers to any form of mistreatment, maltreatment, or harm experienced by children in the school environment, whether it be physical, emotional, or verbal abuse. Such abuse can take severe and lifelong effects on children's emotional well-being, and it is crucial to identify and address these issues to provide appropriate support and interventions for affected children.

Sentiment analysis is the task of categorizing a text as having positive or negative meanings. This task can be considered a classification task with binary (positive, negative) or multi (positive, negative, neutral) classes [1] and there is also a multimodal approach used by [2] using the deep neural networks merging graphic analysis and natural language processing. In [3][4] they used some traditional features which select the optimal feature representations which is a critical step in the use of machine learning in text classification.

In this paper, we proposed a sentiment analysis approach for classifying children's emotional responses into four emotions: Sad, Happy, Angry, and Surprised. Our methodology consists of data pre-processing, feature extraction, classification, and evaluation steps. We plan to apply this approach to a large-scale dataset and provide insights into children's emotions and well-being. The study which we have done can be subsidized to the progress of evidence-based involvements and policies that address children's emotional needs.

For the classification of a dataset having four classes happy, sad, angry, and surprised, our proposed approach for text classification involves the process of loading and pre-processing data, tokenizing text, creating datasets and data loaders, loading a pre-trained BERT model, modifying it for classification, training the model, evaluating its performance, and using it for inference on new data. It leverages the power of pre-trained BERT models to perform classification tasks and fine-tunes the model using custom data for specific classification requirements.

## II. SENTIMENTAL ANALYSIS

### A. Unsupervised Sentimental Analysis

VADER is a rule-based sentiment analysis tool that uses a lexicon of words and their valence scores to determine the sentiment of a text. It considers not only individual words but also the context and linguistic rules to analyze the sentiment of the text. VADER takes into account the intensity and polarity of the words and punctuation such as exclamation marks, question marks, and emoticons. The output of VADER is a sentiment score ranging from -1 to +1,

with 0 indicating neutral sentiment. It provides scores for positive, negative, and neutral sentiment, as well as a compound score that reflects the overall sentiment of the text. VADER is widely used in sentiment analysis for social media monitoring, customer response analysis, and additional applications where analyzing sentiment in short text documents is important [7]. It's possible that dividing emotions into positive, negative, and neutral categories doesn't give a whole picture of how kids feel. Based on the sentiment analysis data produced using VADER. This research divided emotions into four separate groups, namely Happy, Surprised, Angry, and Sad, in order to address this issue. This method gave a finer-grained analysis of the emotional content of the children's monologues, enabling a more sophisticated understanding of their feelings. The researchers were able to acquire a deeper understanding of children's emotional experiences by including these extra categories, which can be helpful for a variety of applications including child psychology, education, and mental health.

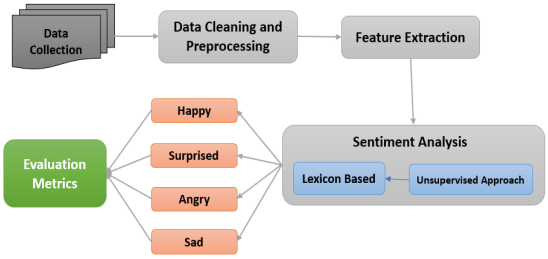



**Fig 1-** Flow Diagram of the Proposed Technique using four classes (Happy, Surprised, Angry, and Sad).

## *B. Supervised Sentimental Analysis*

In this approach, we have datasets consists of sentences and labels of four classes after getting the dataset and loading it we have done the pre-processing of the data by splitting it into sentences and labels. We also create separate lists to store the input sentences, attention masks, and labels for both training and testing data. After that, we use the BERT tokenizer to convert the input sentences into tokens that can be fed into the BERT model. The tokenizer also handles important tasks such as adding special tokens for BERT input format, padding sequences to a fixed length and creating attention masks to indicate which tokens are actual words and which are padding tokens. Then create PyTorch datasets and data loaders to load the tokenized data for training and testing the BERT model. The datasets store the input sentences, attention masks, and labels, while the data loaders provide an efficient way to iterate through the data in batches during model training and evaluation. Now load a pre-trained BERT model from the transformer's library. BERT is an advanced natural language processing (NLP) model that has been pre-trained on a bulky corpus of text data. BERT has a deep architecture with multiple transformer layers and can capture both contextual and positional information from the text and then modify the pre-trained BERT model for classification by adding a classification layer on top of the BERT model. The classification layer is a linear layer with the number of output components equal to the number of classes in the classification task which in our case has four classes (Happy, sad, surprised, and angry). We also define the loss function and optimizer for model training. And trained the modified BERT model using the training dataset. We iterate through the training data in batches, pass the input sentences through the BERT model, calculate the loss, and update the model's weights using backpropagation and the optimizer. This process is repeated for multiple epochs to fine-tune the model on the training data.

At last, we evaluate the trained BERT model using the testing dataset. We pass the input sentences through the trained BERT model, calculate the predicted labels, and compare them with the labels (ground truth) to estimate the accuracy of the model.

And we use the trained BERT model for inference on new data. We encode the new input sentences using the BERT tokenizer, pass them through the trained BERT model, and obtain the predicted labels. We then map the predicted labels back to the original labels and print the predicted labels for the new sentences. Below fig-2 is the proposed flow diagram of the text classification.

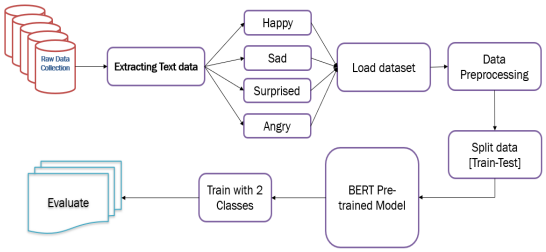



**Fig 2-** Flow Diagram of the Proposed Technique using four classes (Happy, Surprised, Angry, and Sad).

## III. Results of Sentimental Analysis

To accomplish this task, we employ natural language processing techniques, including the bag-of-the-words approach, to extract emotions from the text data. In the first step, we collect data in the form of video and audio files from YouTube searches for 'abused child monologue', convert speech to text, clean the data, remove stop words, punctuation marks, and other extraneous text to convert the text data. It needs to be pre-processed. Then, after a feature extraction process is implemented, the text is tokenized into individual words.

## *A. VADER*

In the VADER approach, we know as per [7] that they have also given the sentiment intensity to the range of -4 to +4 so in our case If the compound score (which represents the overall sentiment) is larger than or equivalent to 0.05, the method classifies the sentence as 'happy'. If the compound score is fewer than or equal to -0.05, the method classifies the sentence as 'sad'. If the positive score is superior than or equivalent to 0.3, the method classifies the sentence as 'surprised'. If the negative score is greater than or equal to 0.3, the method classifies the sentence as 'angry'.

## *B.*

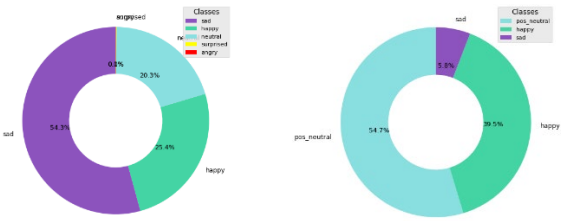


**Fig 3-** Result of classification into four classes (Happy, Surprised, Angry, and Sad). (Abused dataset, left) (non-abused dataset, right).

The chart above is the result of confirming the distribution of emotions through the four classes of VADER's emotion analysis results. The study discovered a distinction between children who had experienced abuse and those who had not. The investigation specifically showed that non-abused children mainly displayed neutral and positive emotions, with melancholy being relatively infrequent, but abused children expressed a wider range of emotions, including happy, negative, and neutral ones. According to these findings, the emotional experiences of abused children are more nuanced and varied, reflecting the psychological effects of abuse on their emotional health.

## *B. BERT*

This segment presents the outcomes of the study based on experimental showed on the dataset having four classes which are in the form of text. The dataset was separated using the ratio of 80:20 into training and testing sets using a suitable method like the train-test split from scikit-learn.. then loading the Bert tokenizer and also pre-trained model and used bert-large-uncased in this proposed method and defining the loss function for multi-class classification tasks which computes the cross-entropy loss between the predicted logits and the ground truth labels. Then to update the weights of the model during training we implement the optimizer AddamW variant of the optimization algorithm and also train the model on 5 epochs. Table I shows the Accuracy, Sensitivity, Specificity, and F1 Score comparison before modification and after modification. In the section before we used "bert-base-uncased" which is a base version of the BERT model, which consists of 12 layers of encoders, and is trained on the uncased text where all the text is converted to lowercase and it has 110 million parameters. And after modifying we have used "bert-large-uncased" which is a larger version of BERT, with 24 layers of an encoder, making it a deeper model and it also has around 340 million parameters. This means that "bert-large-uncased" requires more computational resources for training and inference compared to "bert-base-uncased". Due to its large size and increased depth, "bert-large-uncased" generally tends to perform better than "bert-base-uncased" on NLP tasks.

As is described in Table I name as evaluation metrics we have used the deep learning model BERT to classify the text dataset according to our requirements first we have used the bert-base-uncased and for it we got results which are not that much decent and then we have changed bert-base-uncased to bert-large-uncased (The difference between bert-base-uncased and bert-large-uncased are described above) and also the epochs were changed and increased it after the modification of parameters we have got the promising results as you can see in Table I.

Table I Evaluation Metrix

| | Bert-base-uncased | Bert-large-uncased |
|---|---|---|
| Accuracy | 93.33% | 94.56% |
| Sensitivity | 99.12% | 99.21% |
| Specificity | 86.39% | 99.44% |
| F1 Score | 51.69% | 99.33% |

Below fig 4 shows the confusion matrix which is mapped on the text data after classification and the classification was done on four classes which are explained above in the paper as you can see in the confusion matrix the mapping is according to numeric numbers because the four classes were defined as these four numerical representations in our case '0' considered as 'Happy', '1' considered as 'sad', '2' considered as 'angry' and the fourth class which is named as 'surprised' is considered as '3'.

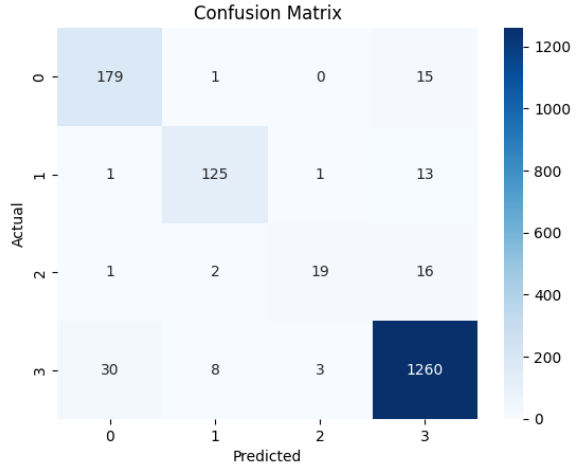



**Fig 4-** Confusion Matrix of Text Classification using four classes with bert-large-uncased.

Our study employed two powerful approaches, VADER and BERT classification, to analyze the sentiment of text data, and both yielded promising results, showcasing their efficacy in sentiment analysis. The VADER approach, known for its simplicity and lexicon-based approach, provided accurate sentiment classification results with minimal computational overhead, making it a practical choice for quick sentiment analysis tasks.

On the other hand, the BERT classification approach, leveraging its deep learning capabilities and contextualized embeddings, demonstrated superior performance in capturing nuanced sentiment patterns, making it ideal for complex text data with varying context and sentiment expressions. While VADER showed strong performance in capturing sentiment polarity, BERT classification excelled in capturing sentiment intensity and fine-grained sentiment nuances, allowing for more granular and nuanced sentiment analysis results.

Our findings reveal that both VADER and BERT classification approaches are powerful tools for sentiment analysis, with VADER being a simple yet effective choice for quick sentiment analysis tasks, and BERT classification being a more sophisticated and accurate option for in-depth sentiment analysis on complex text data. The comparative analysis of VADER and BERT classification demonstrates that both approaches have their strengths and limitations, and the choice of approach depends on the specific requirements of the sentiment analysis task, such as the complexity of the text data, the desired level of sentiment granularity, and the available computational resources. So, If we consider the fewer text data to be labeled then Vader is considered to be the best because it does not contain too much complexity as compared to BERT classification and on another hand if the text data is in more quantity then the better results are of BERT classification

## IV. CONCLUSIONS

In this paper, we proposed a sentiment analysis approach for classifying children's Emotional responses into four emotions: Sad, Happy, Angry, and Surprised. Our methodology consists of data pre-processing, feature extraction, classification, and evaluation steps. We plan to apply this approach to a large-scale survey dataset and provide insights into children's emotions and well-being. Our study can contribute to the development of evidence-based interventions and policies that address children's emotional needs. And for text classification using the deep learning approach, we used four classes Sad, Happy, Angry, and Surprised to classify the text data. In this scenario, our methodology consists of loading and pre-processing the data, splitting data into training and testing datasets, loading and configuring the BERT model, defining the loss function and optimizer, training the BERT model and evaluation steps. And text classification using deep learning methods can be used to predict the nature of the text which would be in sentence form.

## ACKNOWLEDGMENT

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2023-2020-0-01832) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation).

## REFERENCES

[1] Hung, L. P., & Alias, S. (2023). Beyond Sentiment Analysis: A Review of Recent Trends in Text Based Sentiment Analysis and Emotion Detection. Journal of Advanced Computational Intelligence and Intelligent Informatics, 27(1), 84-95.

[2] Anand, M., & Eswari, R. (2019, March). Classification of abusive comments in social media using deep learning. In 2019 3rd international conference on computing methodologies and communication (ICCMC) (pp. 974-977). IEEE.

[3] Chen, H., McKeever, S., & Delany, S. J. (2019, July). The use of deep learning distributed representations in the identification of abusive text. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 13, pp. 125-133).

[4] McKeever, S., & Delany, S. J. (2019). The Use of Deep Learning Distributed Representations in the Identification of Abusive Text.

[5] Gunasekara, I., & Nejadgholi, I. (2018, October). A review of standard text classification practices for multi-label toxicity identification of online content. In Proceedings of the 2nd workshop on abusive language online (ALW2) (pp. 21-25).

[6] Sultana, S., Redoy, M. O. F., Al Nahian, J., Masum, A. K. M., & Abujar, S. (2023). Detection of Abusive Bengali Comments for Mixed Social Media Data Using Machine Learning.

[7] Hutto, Clayton, and Eric Gilbert. "Vader: A parsimonious rule-based model for sentiment analysis of social media text." In Proceedings of the international AAAI conference on web and social media, vol. 8, no. 1, pp. 216-225. 2014.

# Improved Hybrid Swarm Intelligence for Optimizing the Energy in WSN

P.Kavitha Rani[1], JinHyung Kim[2], Yunyoung Nam[3,*]and Mohamed Abouhawwash[4,5]
[1]*Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore,641008, India.*
[2]*B LIFE Inc., Republic of Korea*
[3]*Department of ICT Convergence, Soonchunhyang University, South Korea.*
[4]*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt*
[5]*Department of Computational Mathematics, Science, and Engineering (CMSE), Michigan State University, East Lansing, MI, 48824 USA.*
**Contact: Yunyoung Nam (ynam@sch.ac.kr)*

*Abstract*— **In this current century, most industries are moving towards automation, where human intervention is dramatically reduced. This revolution leads to industrial revolution 4.0, which uses the Internet of Things (IoT) and wireless sensor networks (WSN). With its associated applications, this IoT device is used to compute the received WSN data from devices and transfer it to remote locations for assistance. In general, WSNs, the gateways are a long distance from the base station (BS) and are communicated through the gateways nearer to the BS. At the gateway, which is closer to the BS, energy drains faster because of the heavy load, which leads to energy issues around the BS. Since the sensors are battery-operated, either replacement or recharging of those sensor node batteries is not possible after it is deployed to their corresponding areas. In that situation, energy plays a vital role in sensor survival. Concerning reducing the network energy consumption and increasing the network lifetime, this paper proposed an efficient cluster head selection using Improved Social spider Optimization with a Rough Set (ISSRS) and routing path selection to reduce the network load using Improved Grey wolf optimization (IGWO) approach. (i) Using ISSRS, the initial clusters are formed with the local nodes, and the cluster head is chosen. (ii) Load balancing through routing path selection using IGWO. The simulation results prove that the proposed optimization-based approaches efficiently reduce the energy through load balancing compared to existing systems in terms of energy efficiency, packet delivery ratio, network throughput, and packet loss percentage.**

## I. INTRODUCTION

Wireless Sensors Network (WSN) is a collection of network devices installed in the sensing environment [1]. WSN was widely used in monitoring environments, health care, surveillance, disaster monitoring, and so on [3]. In the network, the individual devices are termed sensor nodes (SN) [2]. These sensor nodes can collect the information from the environment and transfer it into the base station (BS) directly or through intermediate gateways via wireless transmission. The transmission from the SN directly to BS causes an overhead of communication. The sensor nodes are combined as various cluster groups. One among the cluster nodes acts as a head and represents the terminal between SN and BS. This superior node is termed a cluster head (CH). The sensed data is collected from SN within its cluster and transmitted into BS. This kind of WSN is termed a cluster-based WSN.

To improve network performances, WSN has been widely used in various domains [4, 5]. The primary reasons behind using sensors in the environment are their easy configuration and management setup [6]. These sensor nodes are autonomously operated and compute their network infrastructure ad hoc. In this kind of scenario, the nodes are unstable, and they can connect to their Neighbour based on certain factors for data transmission. The CH combines the sensed data and relays it to the BS. CH can efficiently construct its single-hop and multi-hop path toward BS. The end user can use the centralized base station through the internet to retrieve the required data. During data transmission, the deployed SN can be mobile or static. The static SN is referred to as non-adaptive, and its routing mechanism is unchanged. Various mobile SN are dynamic, and their routing tables are updated frequently when there is a change in network topology. While compared to dynamic routing, the static mechanism is secure; moreover, the stationary solutions are not appropriate for the scalable network in more significant regions [7]. The communication efficiency is increased through load balance, resource utilization, and throughput by integrating the Internet of Things (IoT) in all fields [8-10]. More physical devices are involved in IoT to transmit the data using the internet. The WSN helps IoT and supports observing and forwarding the data in the physical networking environment [11]. This research article proposed an effective cluster head selection and routing approach with load balancing to reduce network energy and computation. The significant contribution of this research is as follows:

For Each cluster, the cluster head is selected using Improved Social Spider Optimization with a Roughset model (ISSRS)

- Once the cluster head for each cluster is selected, the routing from source to destination is selected using IGWO

which is used as a load balancing approach to reduce the load of base station and gateways.

- The proposed cluster head selection and load balancing are evaluated using simulation by calculating the total number of ko nodes that are dead and alive. Further energy efficiency, throughput, and execution cost are calculated.

The paper is arranged as follows: Improving the energy efficiency of WSN using existing techniques and methods with load balancing and routing is discussed in section 2. The proposed cluster head selection and routing with load balance methods are introduced in section 3. the simulation environment with evaluation and comparative analysis of proposed Vs. Existing load balancing and routing approaches are discussed in Section 4. Finally, Section 5 concludes the paper with its future extension.

## II. RELATED WORK

The section discusses the existing literature on loading balancing, routing, and clustering in WSN-based IoT systems. Majid et al., [12] did a systematic review of WSN and IoT-based applications. The detailed discussion using WSN in industry 4.0 with its IoT-related applications, research gaps, and future research directions are discussed. Lipare et al., [13] developed GWO with a novel fitness function-based load-balancing approach to overcome the issues of energy holes in WSN. In this work, authors applied GWO for routing and cluster with enhanced fitness functions. Kumar et al. [14] developed a clustering with a Centroid-oriented routing protocol for WSN-based IoT. This routing approach has three phases: network initiation to create the zone among the nodes, selection of zone coordinator, and zone head selection. The path from the zone head to the base station is optimized and distance centric with one or dual-hop level transmission to reduce packet loss.

Kashif et al. [15] used a gateway node to reduce the load of the cluster head and transmit the data to the base station. Their assumption that the gateway node should be in the Neighbour gateway range leads to high complexity and also limits the sensor node deployment randomly. Arroyo et al., [16] developed a cloud computing and WSN-based air quality monitoring system. The distributed sensors are connected to the cloud for storage, monitoring, and processing from WSN. The data processing is performed using artificial intelligence approaches. Mellado et al., [17] predict the shadow of satellites using WSN and IoT for smart cities. The authors developed a graphical method for predicting satellite coverage in urban areas. Their captured shadow images are similar to Eutelsat satellite measurements.

Haseeb et al. [18] proposed a secure IoT with a WSN-based innovative agriculture application. Initially, the agricultural data are sensed using the sensors, and cluster heads are formed using the multi-criteria decision method. The signal strength is measured using the Signal Noise Ratio (SNR). The simulated results prove that the proposed model enhances the network performance with increased throughput and reduced drop ratio. Butun et al., [19] reviewed the challenges and vulnerabilities of IoT and WSN-based applications. Further authors also discussed attacks related to IoT and WSN with its solutions and advantages with limitations.

Pamarthi et al., [20] did a systematic review of the WSN security issues for IoT applications. Pamarthi et al. discussed routing attacks with their security measures for ad hoc networks. Parvathy [21] discussed the wormhole attacks related to WSN and IoT. The authors examined the interconnection between IoT and WSN through the internet and it is guarded against external aggression. Raza et al., [22] reviewed the applications and security issues for sensor networks using the MAC protocol. The security challenges of various layers with their respective solutions were presented.

The technological security issues, with their advantages and challenges, are profoundly concerning. Kumar et al. [23] discussed IoT applications with evolutionary approaches. Sharma et al., [24] designed a relay node structure or WSN and load balancing. Here WSN created an IoT with a WSN-based application to monitor the natural disaster in the coal mine. WSN uses different parameters for network analysis, including the network's lifetime, energy consumption, throughput, and area coverage.

Faheem et al. [25] developed an approach using bio-inspired for WSN-assisted innovative application of grid. This application developed an IoT, and WSN-assisted farming robot using a computer vision-based approach is introduced. It has been used to classify weed and non-image. The weed image data are detected through WSN, and mutual authentication protocols for WSN-based IoT scenarios are also done. This scheme has improved performance, low communication, and execution costs. The main issue is the security concern for heterogeneous IoT environments needs to be considered.

## III. PROPOSED METHODOLOGY

This section proposed efficient cluster head selection and routing path selection for load balancing in WSN assisted IoT environment.

### *A. System Model*

The proposed system overview is shown in Fig. 1. The sensor node (SN) collects the data from the IoT sensors, such as temperature, pressure, and so on, concerning the applications. The gathered data are converted into digital form, which is transferred into the base station (BS) through cluster heads in the wireless networks. The BS consists of more power and memory. Further, it is connected to the best resources to save energy. In IoT-based applications, BS can store the data, analyze it and visualize data collected from the cluster head. BS can also provide a graphical user interface for the user to direct interaction or forward the collected data to the remotely managed server through the internet. The remoter servers are responsible for sending the sensed input data to the authorized user. This data is also saved on web pages to access worldwide through the internet. In this work, the clusters are formed using Euclidean distance, and the cluster head for each group is selected using the ISSRS approach. In the host computer, the load is balanced using IGWO to reduce system energy. The end users can access the centralized base station through the internet to retrieve the required data.

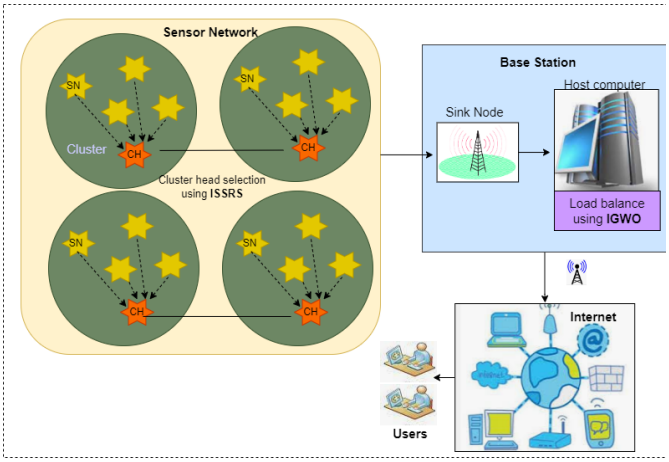



Fig 1. Architecture of proposed load balancing in WSN-assisted IoT model

### B. Cluster Formation

Each SN is allotted its identification number. The links are formed from the source SN to the destination using Euclidean distance. It measures the distance between two adjacent SN located within the communication range CR. The sensing range SNR has been denoted as the range between the SN that can sense the signal. N is the number of nodes in the network. The set of possible communication is defined in Eq. (1).

$$C = \{(SN_i, SN_j) \in N | SN_j \; recieve \; message \; from \; SN_i\} \quad (1)$$

The Neighbouring of $SN_i$ is declared as in Eq. (2)

$$Neighbor(SN_i) = \{SN_j \in N | (SN_i, SN_j) \in l\} \quad (2)$$

The link between the SN is limited to its capacity called $\mu_{ij}$

$$P_{ij} = \sum_{C} X_{ij}^{C} \leq \mu_{ij} \quad (3)$$

Where between $X_{ij}^{C}$ set of communication link between $(i, j) \in L$.

### C. Cluster Head Selection Using ISSRS

Various researchers have proposed cluster-based concepts to simplify network management. The cluster head node can able to manage the sensor node. It is responsible for arranging the SN within the cluster from the routing table. Also, the cluster head gathers the data and aggregates and retransfers the collected data to the collections. These cluster heads lack energy due to the unexpected load assigned to them. To minimize the energy in CH, the distance between CH and the base station should be minimum, or else there is a need for multihop transfer from CH to the base station through intermediate cluster heads. The previous cluster method selects the cluster head node randomly without considering the parameters related to the centralized approach. This will affect the network's scalability. The proposed meta-heuristic-based load balancing and cluster head selection approach efficiently chose the CH to transmit the data to neighbor CH based on the parameters such as Initial energy, Residual energy, Energy consumption rate, and Average network energy. The significant parameter in choosing the CH is the initial energy. Residual energy means after the series of execution, and the CH considers the remaining SN energy. The energy consumption rate is the primary factor for each cluster. The average energy of the network is the reference energy, and it is the ideal energy to make the network alive. This starts with the random initialization of the spider population. Each individual is converted into the binary vector of length N using Eq. (4)

$$V = \frac{1}{1 + e^{-X_i^j(t)}} \quad (4)$$

$$X_i^j(t+1) = \begin{cases} 0 \;\; if \; V\left(X_i^j(t)\right) > \varepsilon \\ 1 \qquad otherwise \end{cases} \quad (5)$$

Where , $X_i^j(t)$- Spider (SN) value at iteration t and $\varepsilon \in [0, 1]$. It uses the dependency degree as in Eq. (6)

$$\gamma_E(D) = \frac{P_E(D)}{|N|} \quad (6)$$

Where, E is the set of conditions and D is the selected nodes, P (D) is the positive region which consists of all the nodes N that is classified based on the information C. The fitness function is declared as in Eq. (7)

$$F\,(S) = \tau\gamma_S(D) + (1-\tau)(1 - \frac{|S|}{|C|}) \quad (7)$$

Where, $\tau \in [0,1]$ provides the distance between number of selected nodes and classifying results. |.| represents the number of selected nodes. Each spider fitness is compared with global best ($F_{gbest}$). If it has highest fitness value then $F_{gbest}$ is replaced with the recent spider and its position which will added to the selected set S. the weight of each spider is updated using the Eq. (8)

$$w_{ij} = \frac{F_i - worst}{best - worst} \quad (8)$$

Where, F – fitness of ith spider, worst –worst fitness value of population N and best is the best fitness value of population N. In social spider the information exchange between the spiders is represented by vibration of ith spider perceived from jth spider is represented in Eq. (9)

$$vb_{i,j} = w_j e^{d_{i,j}^2} \quad (9)$$

Where, $w_j$ th spider weight, *d* distance measured between I and jth spider. There are three categories of vibration produced by

the spider such as n,b, and f. where vbn is the vibration produced by the nearest spider, vbf is the vibration produced by the female spider, and vb is the vibration produced by the male spider as the best spider. The new position of Female Spider Position (FSP) is updated using Eq. (10), which is managed by the probability factor P, and the moving factor is updated with the relation between other spider and their respective vibration in the search space.

$$FSP_i^{k+1} = \begin{cases} FSP_i^k + \alpha.vb_{i,n}.(s_n - FSP_i^k) + \beta.vb_{i,b}.(s_b - FSP_i^k) + \delta.\left(r - \frac{1}{2}\right) with\ probablity\ P \\ FSP_i^k - \alpha.vb_{i,n}.(s_n - FSP_i^k) - \beta.vb_{i,b}.(s_b - FSP_i^k) + \delta.\left(r - \frac{1}{2}\right) with\ probablity\ 1-P \end{cases} \quad (10)$$

Where, $\alpha, \beta, \delta\ and\ r$ - random number between [0,1], k-iteration number and $s_n$and $s_b$ nearest spider and best spider in the web respectively. The male spider process is operated using the Eq. (11). Until the stopping condition met, this process is repeated. The final selected cluster heads are represented in S. The workflow of ISSRS is show in Fig. 2.

$$MSP_i^{k+1} = \begin{cases} MSP_i^k + \alpha.vb_{i,f}.(s_f - MSP_i^k) + \delta.\left(r - \frac{1}{2}\right)\ if MSP_i^k \in Dominant\ group \\ SP_i^k + \alpha.\left(\frac{\sum_{h \in nondominant} MSP_h^k.w_h}{\sum_{h \in nondominant} w_h}\right) - MSP_i^k\ \ if\ MSP_i^k \in Non-Dominant\ group \end{cases} \quad (11)$$

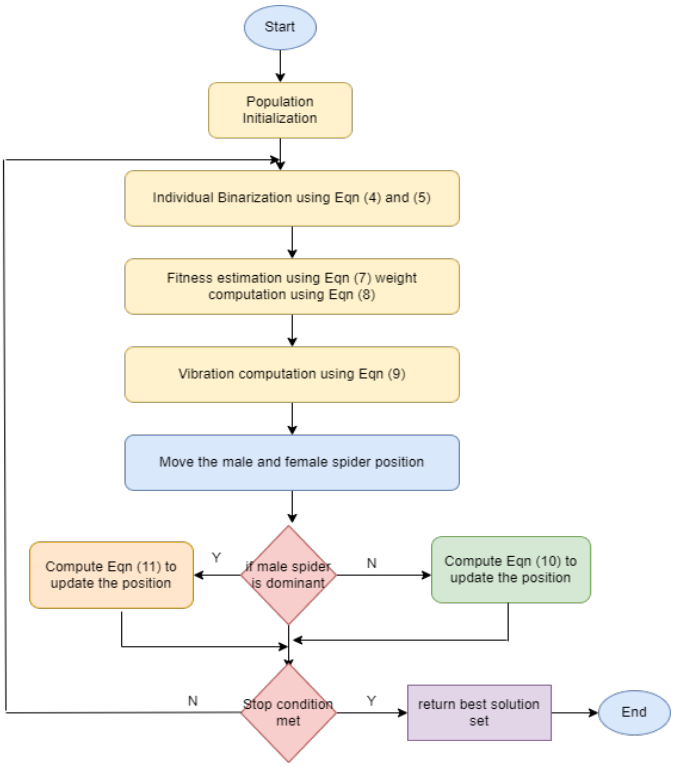



Fig 2. Workflow of ISSRS

## IV. Simulation Results and Discussions

The proposed system model performance is compared with standard Grey wolf optimization, Multi objective fuzzy clustering algorithm (MOFCA), and Particle Swarm Optimization (PSO) - ECHS [26]. The evaluations regarding the number of dead and alive nodes, load balancing, energy consumption, and packet loss are implemented using Python Scikit learn [27-34].

MOFCA approach is made by using both clustering and routing for multi-hop transmission. The proposed approach uses two methods for its efficient routing mechanism. The cluster head selection is based on ISSRS, and the routing is based on IGWO. PSO used only one best possible solution. At the same time, the proposed model used three possible explanations. The routing fitness value comparison of IGWO and other approaches is shown in Tab. 2 for equal and unequal numbers of nodes. From Tab.1. the proposed ISSRS-IGWO secured improved routing fitness values compared to different methods for the equal and unequal load. Increasing the fitness values gives a better solution.

**TABLE I.** ROUTING FITNESS VALUES OF PROPOSED ISSRS-IGWO VS EXISTING MODELS

| Load of SN | Equal loads | | | Unequal loads | | |
|---|---|---|---|---|---|---|
| No. of SN | 100 | 300 | 500 | 100 | 300 | 500 |
| GWO | 1.31 | 1.46 | 1.71 | 0.81 | 0.92 | 0.11 |
| MOFCA | 1.52 | 1.68 | 1.86 | 1.14 | 1.34 | 1.64 |
| PSO-ECHS | 1.42 | 1.53 | 1.61 | 0.91 | 1.2 | 1.7 |
| Proposed ISSRS-IGWO | 1.72 | 1.81 | 1.98 | 1.45 | 1.52 | 1.79 |

Evaluation of the number of alive nodes during the transmission is shown in Fig. 5 for the number of rounds. From the illustration, it has been observed that the proposed model secured longer network life than other approaches. For 3000 matches, using the proposed ISSRS-IGWO model,850 nodes are alive for transmission. At the same time, in the different approaches, such as GWO, MOFCA, and PSO-ECHS, the number of active nodes is 630, 580, and 610, respectively. Hence, the proposed model increases the network lifespan while transmitting the data.

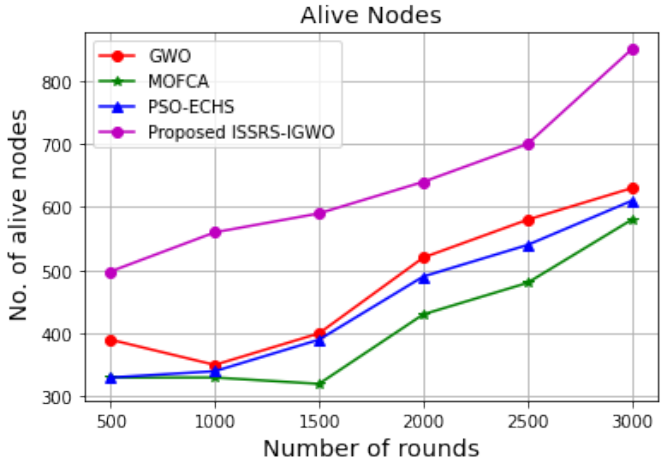



Fig 3. Number of Alive nodes for proposed ISSRS-IGWO vs existing approaches

## V. Conclusions

This paper proposes an efficient cluster head selection and load-balancing approach using Swarm Intelligence for a WSN-assisted IoT environment. Initially, the cluster of sensor nodes is formed using the Euclidean distance between the Neighbour nodes in the communication range. Using the proposed Improved social spider optimization with a rough set approach,

the CH is selected for each cluster, representing the manager for each group. The data from the sensor nodes are forwarded to the base station through cluster heads. The set heads nearer to the base station are sent the sensed data directly. Instead, the clusters far from the base station drain their energy and need extra computation costs to send the data. To reduce the system energy, load, and computation cost, another swarm approach called improved grey wolf optimization had been proposed for efficient routing from SN to BS via CH. The proposed cluster head selection and load balancing were evaluated in the simulation environment regarding the number of dead and alive nodes, energy efficiency, throughput, packet loss, and execution cost. These simulation results are compared with the existing approaches. The results prove the efficiency of the proposed model with reduced energy and execution costs, which will improve the network lifetime. The sensor node mobility is not considered in this work which is regarded as the limitation. In the future, sensor nodes' mobility, cluster head, and base station can be assumed to determine the location change and provides efficient routing with reduced load and energy.

## ACKNOWLEDGMENT

This work was supported by the Collabo R&D between Industry, Academy, and Research Institute(S3250534) funded by the Ministry of SMEs and Startups (MSS, Korea), and the Soonchunhyang University Research Fund.

## REFERENCES

[1] J.Li, L.L.Andrew, C.H.Foh, M.Zukerman and H.H.Chen, "Connectivity coverage and placement in wireless sensor networks,"*Sensors*, vol.9,no.10, pp.7664-7693,2009.

[2] P.Kuila and P.K.Jana, "Improved load balanced clustering algorithm for wireless sensor networks,"in *Proc.International conference on advanced computing, networking and security,* Surathkal, India, pp. 399-404,2011.

[3] J.Tang, B.Hao and A.Sen "Relay node placement in large scale wireless sensor networks,"*Computer Communications*, vol.29,no.4, pp.490-501,2006.

[4] P.S.Mehra, M.N.Doja and B.Alam, "Fuzzy based enhanced cluster head selection (FBECS) for WSN,"*Journal of King Saud University Science*, vol.32,no.1, pp.390-401,2020.

[5] A.Tripathi, H.P.Gupta, T.Dutta, R.Mishra, K.K.Shukla *et al*.,"Coverage and connectivity in WSNs: A survey, research issues and challenges,"*IEEE Access*, vol.6, no. 22, pp.26971-26992,2018,

[6] K.A.Awan, I.U.Din, A.Almogren, M.Guizani and S.Khan, "Stab Trust A stable and centralized trustbased clustering mechanism for IoT enabled vehicular ad-hoc networks," *IEEE Access*, vol.8, no. 4, pp.21159-21177,2020.

[7] A.Abuarqoub,M.Hammoudeh, B.Adebisi, S.Jabbar, A.Bounceur *et al*., "Dynamic clustering and management of mobile wireless sensor networks,"*Computer Networks*, vol.117, no. 11, pp.62-75,2017.

[8] H.A.Khattak, Z.Ameer, U.I.Din and M.K.Khan, "Cross-layer design and optimization techniques in wireless multimedia sensor networks for smart cities,"*Computer Science and Information Systems*, vol.16,no.1, pp.1-17,2019.

[9] I.U.Din,M.Guizani,S.Hassan,B.S.Kim,M.K.Khan *et al*.,"The internet of things: A review of enabled technologies and future challenges," *IEEE Access*, vol.7, no. 4, pp.7606-7640,2018.

[10] K.A.Awan, I.U.Din, A.Almogren, M.Guizani, A.Altameem *et al*.,"Robust trust pro-privacy robust distributed trust management mechanism for internet of things,"*IEEE Access*, vol.7, no. 9, pp.62095-62106,2019.

[11] I.U.Din, M.Guizani, J.J.Rodrigues, S.Hassan and V.V.Korotaev, "Machine learning in the internet of things: Designed techniques for smart cities,"*Future Generation Computer Systems*, vol.100, no.5, pp.826-843,2019.

[12] M.Majid, S.Habib, A.R.Javed, M.Rizwan, G.Srivastava *et al*.,"Applications of wireless sensor networks and internet of things frameworks in the industry revolution 4.0: A systematic literature review,"*Sensors*, vol.22,no.6, pp.2087, 2022.

[13] A.Lipare, D.R.Edla and V.Kuppil, "Energy efficient load balancing approach for avoiding energy hole problem in WSN using grey wolf optimizer with novel fitness function,"*Applied Soft Computing*, vol.84, no. 3, pp.105706,2019.

[14] N.P.R. Kumar and J.B. Gnanadhas, "Cluster centroid based energy efficient routing protocol for WSN assisted IoT," *Advances in Science, Technology and Engineering Systems Journal*, vol.5, no.4, pp. 296-313,2020.

[15] K.N.Qureshi, M.U.Bashir, J.Lloret and A.Le, "Optimized clusterbased dynamic energy aware routing protocol for wireless sensor networks in agriculture precision,"*Journal of Sensors*, vol.2020, no. 4, pp.1-19,2020.

[16] P.Arroyo, J.L.Herrero, J.L.Suárez and J.Loza,"Wireless sensor network combined with cloud computing for air quality monitoring,"*Sensors*, vol.19,no.3, pp. 25-45, 2019.

[17] S.Hornillo,R.Martín and V.Baena, "Prediction of satellite shadowing in smart cities with application to IoT,"*Sensors*, vol.20,no.2, pp.475-487,2020.

[18] K.Haseeb, I.Ud Din, A.Almogren and N.Islam, "An energy efficient and secure IoT-based WSN framework: An application to smart agriculture,"*Sensors*, vol.20,no.7, pp.2081-2095,2020.

[19] I.Butun,P.Österberg and H.So, "Security of the Internet of Things: Vulnerabilities, attacks, and countermeasures,"*IEEE Communications Surveys & Tutorials*, vol.22,no.1, pp.616-644,2019.

[20] S.Pamarthi and R.Narmadha, "Literature review on network security in wireless mobile ad-hoc network for IoT applications: Network attacks and detection mechanisms,"*International Journal of Intelligent Unmanned Systems*, vol. 6, no. 2, pp. 1-22, 2021.

[21] M.U.Farooq,M.Waseem, S.Mazhar,A.Khairi and T.Kamal, "A review on internet of things (IoT),"*International Journal of Computer Applications*, vol.113,no.1, pp.1-7,2015.

[22] S.Raza, M.Faheem and M.Guenes, "Industrial wireless sensor and actuator networks in industry 4.0: Exploring requirements, protocols, and challenges—A MAC survey,"*International Journal of Communication Systems*, vol.32,no.15, pp.e4074,2019.

[23] S.Kumar, P.Tiwari and M.Zymbler, "Internet of Things is a revolutionary approach for future technology enhancement: a review,"*Journal of Big Data*, vol.6,no.1, pp.1-21,2019.

[24] R.Sharma and S.Prakash, "Enhancement of relay nodes communication approach in WSNIoT for underground coal mine,"*Journal of Information and Optimization Sciences*, vol.41,no.2, pp.521-531,2020.

[25] M.Faheem, R.A.Butt, B.Raza, M.W.Ashraf and S.Begum, "Bio-inspired routing protocol for WSN-based smart grid applications in the context of Industry 4.0,"*Transactions on Emerging Telecommunications Technologies*, vol.30,no.8, pp.e3503,2019.

[26] P.S.Rao, P.K.Jana and H.Banka, “A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks,”*Wireless Network*,vol.23, no. 4, pp.2005–2020,2017.

[27] A.H. El.Bassiouny, M. Abouhawwash and H.S. Shahen, "New generalized extreme value distribution and its bivariate extension," *International Journal of Computer Applications*, vol. 173, no. 3, pp. 1-10, 2017.

[28] A.H.El.Bassiouny, M.Abouhawwash and H.S.Shahen, "Inverted exponentiated gamma and it’s a bivariate extension," *International Journal of Computer Application*, vol. 3, no. 8, pp. 13-39, 2018.

[29] M. Abouhawwash and M.A. Jameel, "KKT proximity measure versus augmented achievement scalarization function," *International Journal of Computer Applications*, vol. 182, no. 24, pp. 1-7, 2018.

[30] H.S.Shahen, A.H.El.Bassiouny and M.Abouhawwash, "Bivariate exponentiated modified Weibull distribution," *Journal of Statistics Applications & Probability*, vol. 8, no. 1, pp. 27-39, 2019.

[31] M. Abouhawwash and M.A.Jameel, "Evolutionary multi-objective optimization using Benson’skush-Kuhn-tucker proximity measure," in *Proc.International Conference on Evolutionary Multi-Criterion Optimization*, East Lansing, Michigan, USA, pp. 27-38, 2019.

[32] M.Abdel Basset, R. Mohamed, M. Elhoseny, M. Abouhawwash, Y. Nam *et al*., "Efficient MCDM model for evaluating the performance of commercial banks: A case study," *Computers, Materials & Continua*, vol. 67, no. 3, pp. 2729-2746, 2021.

[33] B.Gomathi, S.Balaji, V.K.Kumar, M.Abouhawwash, S. Aljahdali *et al*., "Multi-objective optimization of energy aware virtual machine placement in the cloud data center," *Intelligent Automation & Soft Computing*, vol.33, no.3, pp. 1771-1785, 2022.

[34] M.Kumar, K.Venkatachalam, M. Masud and M. Abouhawwash, "Novel dynamic scaling algorithm for energy-efficient cloud computing," *Intelligent Automation & Soft Computing*, vol.33, no.3, pp. 1547-1559, 2022.

# Energy-Efficient Clustering Using Optimization with Locust Game Theory

P.Kavitha Rani[1], Hee-Kwon Chae[2], Yunyoung Nam[2,*] and Mohamed Abouhawwash[3,4]

[1]*Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore-641008, India.*

[2]*Department of ICT Convergence, Soonchunhyang University, Asan, 31538, Korea.*

[3]*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.*

[4]*Department of Computational Mathematics, Science, and Engineering (CMSE), Michigan State University, East Lansing, MI, 48824 USA.*

**Contact: Yunyoung Nam (ynam@sch.ac.kr)*

*Abstract*— **Wireless sensor networks (WSNs) are made up of several sensors located in a specific area and powered by a finite amount of energy to gather environmental data. WSNs use sensor nodes (SNs) to collect and transmit data. However, the power supplied by the sensor network is restricted. Thus, SNs must store energy as often as to extend the lifespan of the network. In the proposed study, effective clustering and longer network lifetimes are achieved using multi-swarm optimization (MSO) and game theory based on locust search (LS-II). In this research, MSO is used to improve the optimum routing, while the LS-II approach is employed to specify the number of cluster heads (CHs) and select the best ones. After the CHs are identified, the other sensor components are allocated to the closest CHs to them. A game theory-based energy-efficient clustering approach is applied to WSNs. Here each SN is considered a player in the game. The SN can implement beneficial methods for itself depending on the length of the idle listening time in the active phase and then determine to choose whether or not to rest. The proposed multi-swarm with energy-efficient game theory on locust search (MSGE-LS) efficiently selects CHs, minimizes energy consumption, and improves the lifetime of networks. The findings of this study indicate that the proposed MSGE-LS is an effective method because its result proves that it increases the number of clusters, average energy consumption, lifespan extension, reduction in average packet loss, and end-to-end delay.**

## I. INTRODUCTION

Wireless Sensor Networks (WSNs) are a collection of cost-effective, multipurpose Sensor Nodes (SNs) which work together to detect a specific area as an Area of Interest (AoI). The system collects the data from the AoI and sends it to a Base Station (BS) for processing. WSNs have made it simpler to monitor remote places. They are particularly successful in gathering information in a variety of remote locations, including rescue teams, forests, war-prone regions, undersea research, and weather patterns. WSNs have numerous SNs that are connected to a BS. However, it is difficult to power and replace SNs in remote places [1].

SNs in WSNs help to detect communication and gather data collaboratively. A certain quantity of energy is consumed during this procedure. However, SNs are powered by a rechargeable battery and have a finite quantity of energy in case the batteries die. Also, they do not need to change the battery pack or update it promptly. The communication of the data collected will be hampered to some extent, and the entire sensor network may become incapacitated [2]. Consequently, extending the lifespan of the complete sensor network is energy-efficient and has become a major challenge in real-world remote sensing applications.

In real-world applications, wireless SNs are powered by limited-capacity batteries and WSNs.This node is placed unprotected or in unsafe places in a challenging way. They are powered by batteries, and the surveillance region is unregulated. This makes it difficult to store energy supplies and rechargeable batteries. Thus, decreasing the energy consumption of the network seems to be the most effective method to extend the lifespan of the sensor network [3] and ensure the continuous operation of WSNs. Many professionals and academicians have undertaken various types of research to optimize network energy consumption to a certain extent [4–9].

Cluster-based or hierarchical routing is a common mode of communication with appealing qualities such as efficiency and flexibility. The general concepts of routing protocol have been combined with most of the energy-saving approaches in WSNs. Nodes with lots of energy are evaluated for their abilities to analyze and communicate the data in the framework of a hierarchical technique [10]. Low-energy terminals on either side are utilized to gather data in areas close to the objective. Hierarchical navigation is an effective method to reduce power usage within clustered architectures by aggregating the data. Hierarchical routing approaches combine operations to minimize the number of transmitted packets for delivery to a specific sink [11].

In these circumstances, assigning special roles to nodes can considerably extend the lifespan of the network and the flexibility of design. Cluster Heads (CHs) are special nodes that serve as anchors in the hierarchy or cluster-based systems. Scaling is a key attribute in WSNs. Owing to the limits given by the original assumption, this idea is still a challenge that

hasn't been solved in most of the routing approaches. Cluster-based methods find single sinks and a few CHs in a general WSN to increase the influence of their coverage spaces [12].

The assignment of unique purposes to nodes in such circumstances can considerably extend the lifetime of the system and the adaptability of technology. CHs are the special nodes that serve as referrals in hierarchical or clustered systems. The Locust Search (LS-II) method [13] is a meta-heuristic optimization technique derived from the modeling of desert locust biological behavior. A locust acts in two ways; alone and in groups. The solitary method prevents the components from interacting with other entities to discover new sources of food. The sociological perspective, on either hand, takes into account the large proportion of entities in locations with ample food supplies.

The locust concentrating process includes the movement of an entity in lowland areas to all those components in locations with the greatest food supplies. The LS-II technique combines both tendencies to produce highly powerful global and regional searching capabilities.

These properties have led to the application of LS-II in a wide range of challenging optimization formulas, including parameterization in chaotic systems [14], image processing [15–16], and pattern recognition [17].

The main contributions of the proposed method are given below:

- Identify the CHs effectively and minimize energy consumption using the LS-II algorithm
- Multi-Swarm Optimization (MSO) helps to find the optimal routing path, and the SN energy is stored feasibly.
- A penalty technique is designed to motivate SNs to embrace collaborative tactics in operational plans to prevent their destructive behavior if they go to sleep.
- For sensor networks exhibiting selfish behavior, the optimum amount of penalty sessions are established to efficiently improve the lifetime of the network.

The remaining sections of this paper are structured as follows; Section 2 discusses the related research. Section 3 describes the WSNs and optimization methods for minimizing energy consumption. Section 4 discusses the experimental results and comparison. Section 5concludes the proposed optimization method with future work.

## II. RELATED WORK

Whenever an unbalanced situation occurs in information transfer across the transmission network, an energetic vacuum is generated, causing the sink nodes to die prematurely and reducing the lifespan of the network [18,19]. Researchers [20,21] offered a strategy wherein the communication range was changed depending on the proximity between the CHs and their users. The use of the firefly optimization technique also extended the longevity of the program. The method was applied against the standardized methods in a variety of scenarios. According to the findings of their research, the proposed method produced good outcomes in terms of network life. The author, on the other hand, made no mention of the energy maximization of SNs while preserving the optimal CH range.

The researcher [22] proposed the particle swarm optimization-based unequal fault-tolerant clustering (PSO-UFC) technology. Unbalanced grouping and fault-tolerant difficulties of the present energy-balanced unequal clustering technique were solved effectively in the method for the long-term network. An absence of inequity in the clustering method was used to equalize intra-cluster and inter-cluster power use among the master CHs (MCHs) to find a remedy for the unbalanced in the clustering algorithm. Moreover, with the PSO-UFC method, the internet connection was restored by selecting an additional CH, termed substitute CH, in the event of an impetuous MCH breakdown. The lifespan maximization of SNs, on the other hand, has appeared to be dependent on an optimized clustering method that reduces end-to-end delay.

In [23], the author proposed regional energy-aware clustering with isolated nodes (REAC-IN) for grouping. The CH was chosen in REAC-IN based on the weight, and the value was calculated using the residual energy of each sensor and the average overall energy of all sensors in each cluster. In the end, the proposed approach had good results when compared to the other WSN protocols, although numerous difficulties had been overlooked. A fresh, enhanced method was developed to overcome those limitations. Grouping and task selection were used to demonstrate the suggested technique which could transcend the limits of the REAC-IN routing algorithm. Their suggested methodology surpassed other methodological approaches.

The researchers [24] used an upgraded harmony search (HS)-based routing to transport the incoming packets between CHs and the sinks from another study. The efficiency of the combined grouping and forwarding system was assessed using energy cost, the proportion of live and dead components, and system longevity. In comparison with the existing methods, the novel cuckoo–HS-based combined navigation and grouping method provided more results. When constructing the state-of-the-art meta-heuristic algorithms with merged forwarding and clustering, the researcher disregarded incorporating the lifespan and end-to-end latency elements.

## III. PROPOSED MSGE-LS METHODOLOGY

The proposed multi-swarm with energy-efficient game theory on locust search(MSGE-LS)method consists of two models; the energy model and the network model. The proposed method efficiently selects the CH and optimally chooses the route for routing. It effectively minimizes energy consumption and improves the life span of the network. Fig.1 shows the architecture of the proposed method. MSO with game theory optimally chooses the routing path, and LS-II helps to select the CH successfully.

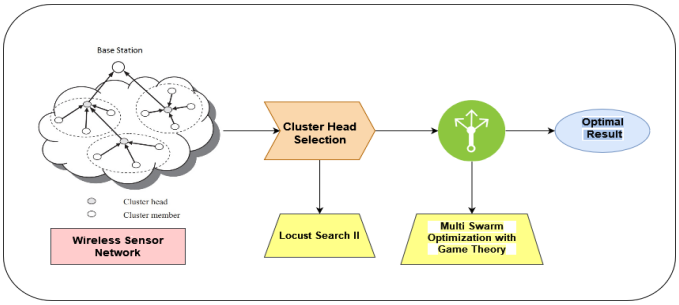



Fig 1. Architecture of the proposed method

### A. Network Model

A WSN is a network made up of a randomized number of nodes that form various groups, each with a CH node and many cluster members (CMs). Every CM has a detection range through which each node can receive information from the observed item and transmit it to the appropriate CH, who subsequently sends it to the sink node.

The functioning phases of an SN (exit phase) occur when the node has to transmit and receive information, and the inactive monitoring phase occurs when a network has no information to transmit and receive. To save power, all functionalities on the SN are switched off when it is in silent mode. The energy usage of the SN can be regarded as zero in this state. In this work, the SNs are considered to be appropriate and the energy in severe conditions is restricted and divided. As an advantage, the SN will be feasible to fulfill the goal of maximizing life span.

Therefore, to save power usage, maximize the lifespan of the network, and ensure the performance of the network, the SN must reach a latent state as much as feasible when it is in the inactive monitoring phase.

### *B. Energy Model*

The energy consumption of the SN is influenced by a variety of variables. The primary elements that have a direct impact on the energy consumption of the SN are listed below. Different energy consumption aspects of SNs throughout the process of information need to be investigated to compute the power leftovers of the SN.

### *C. Locust Search II*

The LS-II approach is a metaheuristic method developed by studying the social behavior of locust swarms. The methods in LS-II prevent individuals from congregating in favorable locations and promoting the search area by redistributing the agents [25]. The LS-II technique combines both the tendencies to produce highly efficient domestic and global search capabilities. These properties have led to its application in solving difficult problems [26–43].

LS-II is a follow-up of the initial LS. Unlike the original LS, LS-II includes additional controllers and processes to boost its performance and reduce the accumulation of agents. Whenever the procedure arrives with multiple locally optimal solutions, such methods allow for a more accurate balance between exploratory and exploitative in identifying the global solution.

A population (PO) of N locusts constitutes a group of N possible solutions in LS-II, that is, PO =$\{ x_1,..., x_N\}$. The parts of PO communicate with one another as they explore an n-dimensional area. Every solution $x_i = [x_{i,1},..., x_{i,n}]$ is located within the limited space UN = {$x_i$ Rn|ld xi,d ud (where $l_d$ and $u_d$ represent the minimum and maximum values for the d-th choice variables, respectively). LS-II, like any other metaheuristic technique, is based on an iterative development wherein the optimal solutions change their position with every repetition. The position of each candidate solution is changed by using several methods after the two behavioral patterns seen in the locust insects, namely, isolated and social activities.

## IV. ANALYSIS OF RESULTS

The suggested MSGE-LS method is tested in this part by evaluating its efficiency with that of GA, LS, and MSO-based clustering. The statistical tests are run on a Dell OptiplexTM 3020 PC featuring an Intel Core i7-4770 3.4 GHz processor, 8 MB cache, inbuilt Intel graphics, 16 GB RAM, 1 TB HDD, and Windows 10 operating system. For the simulations, Matlab R2019a is employed. Tab.1 lists the network parameters that were employed during the process.

TABLE I SIMULATION PARAMETERS

| Parameters | Representations |
|---|---|
| Ranges between transmission | 30 m |
| Sensing distances among nodes | 10 m |
| Initial energy of nodes | 50 J |
| Bandwidth of nodes | 50 Kbps |
| Network's size | 400 sqm |
| Transmission rate of packets | 30 packets/s |
| Time taken for simulation | 32 min |
| Packet's size | 8,16,32,64,128,256 bytes |
| Mobility | 0.5–3.5 m/s |

The experiment of the proposed method is calculated using 200–1400 nodes and 0–800 rounds. The criteria that were taken into account are stated; number of clusters that generated the overall end-to-end delay in seconds, mean packet loss rates, lifespan computations, throughput, and total power dissipated in joules. The experimental calculations are shown in Tab. 2.

TABLE II TOTAL NUMBER OF CLUSTERS GENERATED IN MSGE-LS

| Total number of nodes | GA | LS | MSO | MSGE-LS |
|---|---|---|---|---|
| 200 | 10 | 11 | 13 | 15 |
| 400 | 15 | 16 | 18 | 21 |
| 600 | 24 | 23 | 25 | 29 |
| 800 | 27 | 28 | 27 | 32 |
| 1000 | 29 | 30 | 31 | 36 |
| 1200 | 29 | 31 | 33 | 40 |
| 1400 | 32 | 36 | 39 | 43 |

The proposed MSGE-LS method forms 4% of clusters higher than the other methods in 1400 rounds. The existing GA method forms 32 clusters, LS forms 36 clusters, and MSO forms 39 clusters in 1400 rounds. For 200, 400, 600, 800, 1000, 1200, and 1400 nodes, MSGE-LS improves approximately by 32%, 15%, 10%, 9%, 16%, and 13%, respectively, while compared with GA. MSGE-LS also outperforms LS by 9%, 14, 7%, 5%, 6% and 0. Fig.3 shows the number of clusters formed during 1400 nodes.

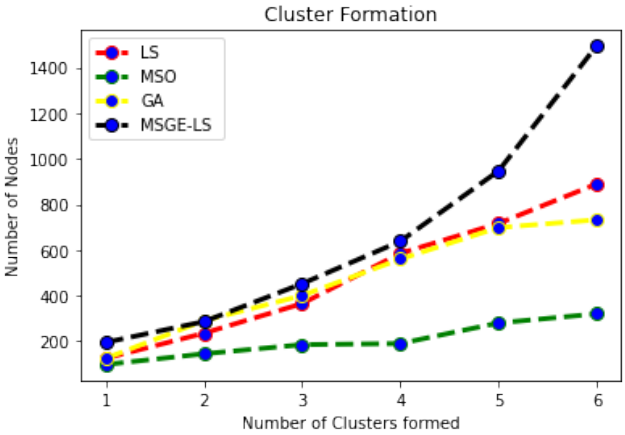



Fig 2. Formations of clusters

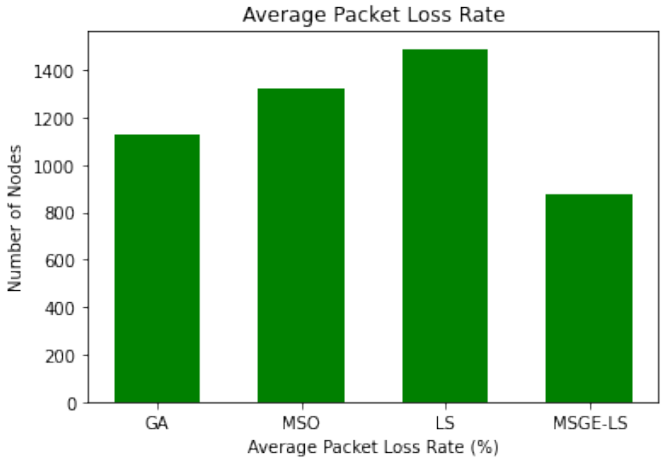



Fig 3. Average packet loss rate

The proposed MSGE-LS method achieves 65.31% of the minimum packet loss rate compared to the existing methods. In the existing methods, GA achieves 84.56%, LS achieves 80.35%, and MSO achieves 88.45% in average packet loss rates for 1400 rounds. For 200, 400, 600, 800, 1000, 1200, and 1400 nodes, the MSGE-LS has a reduced loss rate of 20%, 22%, 24%, 19%, 8%, and 20% when compared with GA and 12%, 11%, 18%, 17%, 11%, 9%, and 19% when compared to LS. MSGE-LS is also found to be more effective than MSO-based clustering by 6%, 5%, 5%, 7%, 4%, and 3% for 200, 400, 600, 800, 1000, and 1200 nodes, correspondingly. The experimental calculations are shown in Tab. 3. Fig.4 shows the average packet loss rate.

TABLE III. AVERAGE PACKET LOSS RATE

| Number of nodes used | GA | LS | MSO | MSGE-LS |
|---|---|---|---|---|
| 200 | 24.06 | 29.04 | 17.95 | 15.12 |
| 400 | 42.51 | 32.6 | 31.63 | 28.15 |
| 600 | 52.06 | 48.19 | 52.18 | 33.84 |
| 800 | 69.51 | 62.09 | 57.64 | 55.41 |
| 1000 | 73.56 | 71.42 | 64.54 | 51.45 |
| 1200 | 76.11 | 79.66 | 75.15 | 63.16 |
| 1400 | 84.56 | 80.35 | 88.45 | 65.31 |

## V. CONCLUSIONS

Two of the most essential criteria for operating WSNs successfully and consistently are energy conservation and energy consumption. At present, more importance is given to technical issues and resource management. The main goal of this research is to use the LS-II method to improve WSN CH selection. Therefore, a method that relies on MSO and game theory has been developed to improve WSN and CH selection routing. The solution generated was based on both energy efficiency and network responsiveness. The suggested method optimizes routing and extends the lifetime of the network by selecting the effective CHs in reaction to the networks. According to the findings, the suggested scheme outperforms the comparative methods in simulations. The suggested protocol is evaluated using several parameters, such as the lifetime of the network, energy usage, energy divergence, number of packets transmitted, and system stability duration. The research has shown that the suggested methodology performs well in terms of increasing the lifespan of the network and providing a good energy load.

The limitation of this research is the lack of energy efficiency across the network. In such a situation, the challenging due nodes are considered as players. If one node acts as a player, the whole algorithm takes responsibility only for that player's node. In the future, a better optimization model can be used for node optimization.

## ACKNOWLEDGMENT

This research was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0012724, The Competency Development Program for Industry Specialist) and the Soonchunhyang University Research Fund.

## REFERENCES

[1] S.Suganthi, N.Umapathi, M.Mahdal and M.Ramachandran, "Multi swarm optimization based clustering with tabu search in wireless sensor network," *Sensors*, vol.22, no.5, pp.1736-1745,2022.

[2] A.Rodríguez, M.P.Cisneros, J.C.R.Caro, C.D.V.Soto, J.Gálvez *et al*., "Robust clustering routing method for wireless sensor

networks considering the Locust search scheme," *Energies*, vol.14,no.11, pp.3019-3025,2021.

[3] X.Yan, C.Huang, J.Gan and X.Wu, "Game theory based energy-efficient clustering algorithm for wireless sensor networks," *Sensors*, vol.22, no.2, pp.478-485,2022.

[4] X.Huan, K.S.Kim, S.Lee, E.G.Lim and A. Marshall, "A beaconless asymmetric energy-efficient time synchronization scheme for resource-constrained multi-hop wireless sensor networks," *IEEE Transactions on Communications*, vol.68, no.3, pp.1716-1730, 2019.

[5] S.Chamanian, S.Baghaee, H.Uluşan, O.Zorlu, E.U. Biyikoglu *et al*., "Implementation of energy-neutral operation on vibration energy harvesting WSN, "*IEEE Sensors Journal*, vol.19, no.8, pp.3092-3099, 2019.

[6] R.Ashween, B.Ramakrishnanand and M.M. Joe, "Energy efficient data gathering technique based on optimal mobile sink node selection for improved network lifetime in wireless sensor network (WSN),"*Wireless Personal Communications*, vol.113, no.4, pp.2107-2126, 2020.

[7] S.Q.Zhang, Y.Tao and J.J.Dai, "Multi-hop clustering routing protocol for energy harvesting wireless sensor networks," *Computer Engineering Design*, vol.40, no. 4, pp.611-616, 2019.

[8] N.Saxena, A.Roy and J.Shin, "Dynamic duty cycle and adaptive contention window based QoS-MAC protocol for wireless multimedia sensor networks," *Computer Networks*, vol.52*,* no.13, pp.2532-2542, 2018.

[9] M.B.Dowlatshahi, M.K.Rafsanjani and B.B.Gupta, "An energy-aware grouping memetic algorithm to schedule the sensing activity in WSNs-based IoT for smart cities," *Applied Soft Computing*, vol.108, no.2, pp.107473, 2021.

[10] S.Kassan, J.Gaber and P.Lorenz, "Game theory based distributed clustering approach to maximize wireless sensors network lifetime," *Journal of Network and Computer Applications*, vol.123, no. 11, pp.80-88, 2018.

[11] S.Shen, H.Ma, E.Fan, K.Hu, S.Yu *et al*., "A non-cooperative non-zero-sum game-based dependability assessment of heterogeneous WSNs with malware diffusion," *Journal of Network and Computer Applications*, vol.91, no.3, pp.26-35, 2017.

[12] A.González, E.Cuevas, F.Fausto, A.Valdivia and R.Rojas, "A template matching approach based on the behavior of swarms of locust," *Applied Intelligence*, vol.47, no.4, pp.1087-1098, 2017.

[13] X.Fu, P.Pace, L.Aloi Yang and G.Fortino, "Topology optimization against cascading failures on wireless sensor networks using a memetic algorithm," *Computer Networks*, vol.177, no.3, pp.107327, 2020.

[14] X.Cheng, D.Ciuonzo and P.S.Rossi, "Multibit decentralized detection through fusing smart and dumb sensors based on Rao test," *IEEE Transactions on Aerospace and Electronic Systems*, vol.56,no.2, pp.1391-1405,2019.

[15] D.Ciuonzo, P.S.Rossi and P.K.Varshney, "Distributed detection in wireless sensor networks under multiplicative fading via generalized score tests," *IEEE Internet of Things*, vol.8,no.11, pp.9059-9071, 2021.

[16] A.Shahraki, A.Taherkordi, O.Haugen and F.Eliassen, "Clustering objectives in wireless sensor networks: A survey and research direction analysis," *Computer Networks*, vol.180, no.4, pp.107376, 2020.

[17] M.Krishnan, S.Yun and Y.M.Jung, "Improved clustering with firefly-optimization-based mobile data collector for wireless sensor networks," *AEU-International Journal of Electronics and Communications*, vol.97, no. 3, pp.242-251, 2018.

[18] M.Krishnan, Y.M.Jung and S.Yun, "An improved clustering with particle swarm optimization-based mobile sink for wireless sensor networks," in *Proc.2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, pp.11-12, 2018.

[19] Y.Xu, O.Ding, R.Qu and K.Li, "Hybrid multi-objective evolutionary algorithms based on decomposition for wireless sensor network coverage optimization," *Applied Soft Computing*, vol.68, no. 4, pp.268-282,2018.

[20] T.Kaur and D.Kumar, "Particle swarm optimization-based unequal and fault tolerant clustering protocol for wireless sensor networks," *IEEE Sensors Journal*, vol.18,no.11, pp.4614-4622,2018.

[21] V.Varsha, J.Singh and M.Bala, "Tabu search based energy efficient clustering protocol for wireless sensor networks," *Global Journal of Computers & Technology*, vol.5,no.2, pp.302-309,2017.

[22] G.P.Gupta, "Improved cuckoo search-based clustering protocol for wireless sensor networks," *Procedia Computer Science*, vol.125, no. 6, pp.234-240, 2018.

[23] L.Kong, J.S.Pan, V.Snášel, P.W.Tsai and T.W.Sung, "An energy-aware routing protocol for wireless sensor network based on genetic algorithm," *Telecommunication Systems*, vol.67, no.3, pp.451-463, 2018.

[24] J.Wang, Y.Gao, W.Liu, A.K.Sangaiah and H.J.Kim, "An improved routing schema with special clustering using PSO algorithm for heterogeneous wireless sensor network," *Sensors*, vol.19,no.3, pp.671-684,2019.

[25] M. Abouhawwash and K. Deb, "Karush-Kuhn-tucker proximity measure for multi-objective optimization based on numerical gradients," in *Proc. 2016 on Genetic and Evolutionary Computation Conference Companion, ACM*, Denver Colorado USA, pp. 525-532, 2016.

[26] A.H. El.Bassiouny, M. Abouhawwash and H.S. Shahen, "New generalized extreme value distribution and its bivariate extension," *International Journal of Computer Applications*, vol. 173, no. 3, pp. 1-10, 2017.

[27] A.H.El.Bassiouny, M.Abouhawwash and H.S.Shahen, "Inverted exponentiated gamma and it's a bivariate extension," *International Journal of Computer Application*, vol. 3, no. 8, pp. 13-39, 2018.

[28] A.H.El.Bassiouny, H.S.Shahen and M.Abouhawwash, "A new bivariate modified Weibull distribution and its extended distribution," *Journal of Statistics Applications & Probability*, vol. 7, no.2, pp. 217-231, 2018.

[29] M. Abouhawwash and M.A. Jameel, "KKT proximity measure versus augmented achievement scalarization function," *International Journal of Computer Applications*, vol. 182, no. 24, pp. 1-7, 2018.

[30] H.S.Shahen, A.H.El.Bassiouny and M.Abouhawwash, "Bivariate exponentiated modified Weibull distribution," *Journal of Statistics Applications & Probability*, vol. 8, no. 1, pp. 27-39, 2019.

[31] M. Abouhawwash and M.A.Jameel, "Evolutionary multi-objective optimization using Benson's kush-Kuhn-tucker proximity measure," in *Proc.International Conference on Evolutionary Multi-Criterion Optimization*, East Lansing, Michigan, USA, pp. 27-38, 2019.

[32] M. Abouhawwash, M.A. Jameel and K. Deb, "A smooth proximity measure for optimality in multi-objective optimization using Benson's method," *Computers& Operations Research*, vol. 117, no. 2, pp. 104900, 2020.

[33] M.Abouhawwash, K.Deb and A. Alessio, "Exploration of multi-objective optimization with genetic algorithms for PET image

reconstruction," *Journal of Nuclear Medicine*, vol. 61, no. 1, pp. 572-572, 2020.

[34] M.AbdelBasset, R. Mohamed, M. Elhoseny, M. Abouhawwash, Y. Nam *et al*., "Efficient MCDM model for evaluating the performance of commercial banks: A case study," *Computers, Materials & Continua*, vol. 67, no. 3, pp. 2729-2746, 2021.

[35] B.Gomathi, S.Balaji, V.K.Kumar, M.Abouhawwash, S. Aljahdali *et al*., "Multi-objective optimization of energy aware virtual machine placement in the cloud data center," *Intelligent Automation & Soft Computing*, vol.33, no.3, pp. 1771-1785, 2022.

[36] M.Kumar, K.Venkatachalam, M. Masud and M. Abouhawwash, "Novel dynamic scaling algorithm for energy-efficient cloud computing," *Intelligent Automation & Soft Computing*, vol.33, no.3, pp. 1547-1559, 2022.

[37] R.S.Ram, K.Venkatachalam, M.Masud and M.Abouhawwash, "Air pollution prediction using dual graph convolution LSTM technique," *Intelligent Automation & Soft Computing*, vol.33, no.3, pp. 1639-1652, 2022.

[38] A.J. Basha, N. Rajkumar, M.A. AlZain, M. Masud and M. Abouhawwash, "Fog based self-sovereign identity with RSA in securing IoMT data," *Intelligent Automation & Soft Computing*, vol. 34, no. 3, pp.1693-1706, 2022.

[39] G. Ravikumar, K. Venkatachalam, M.A. AlZain, M. Masud and M. Abouhawwash, "Neural cryptography with fog computing network for health monitoring using IoMT," *Computer Systems Science and Engineering*, vol. 44, no.1, pp.945-959, 2023.

[40] R. Rajdevi, K. Venkatachalam, M. Masud, M.A. AlZain and M. Abouhawwash, "Proof of activity protocol for IoMT data security," *Computer Systems Science and Engineering*, vol. 44, no. 1, pp.339-350, 2023.

[41] G. Ravikumar, K. Venkatachalam, M. Masud and M. Abouhawwash, "Cost-efficient scheduling using smart contract cognizant Ethereum for IoMT," *Intelligent Automation & Soft Computing*, vol.33, no.2, pp. 865-877, 2022.

[42] N. Mittal, H. Singh, V. Mittal, S. Mahajan, A.K. Pandit *et al*., "Optimization of cognitive radio system using self-learning salp swarm algorithm," *Computers, Materials & Continua*, vol.70, no.2, pp.3821-3835, 2022.

[43] A.Garg, A.Parashar, D.Barman, S.Jain, D.Singhal *et al*., "Autism spectrum disorder prediction by an explainable deep learning approach," *Computers, Materials & Continua*, vol.71, no.1, pp.1459-1471, 2022.

# Improved Shark Smell Optimization Algorithm for Human Action Recognition

Inzamam Mashood Nasir[1], Mudassar Raza[1], Jamal Hussain Shah[1], Muhammad Attique Khan[2], Yun-Cheol Nam[3] and Yunyoung Nam[4, *]

[1]*Department of Computer Science, COMSATS University Islamabad, Wah Campus, Wah Cantt. 47040, Pakistan.*
[2]*Department of Computer Science, HITEC University, Taxila, Pakistan.*
[3]*Department of Architecture, Joongbu University, Goyang, 10279, South Korea.*
[4]*Department of ICT Convergence, Soonchunhyang University, Asan 31538, Korea*
**Contact: Yunyoung Nam (ynam@sch.ac.kr)*

*Abstract*— **Human Action Recognition (HAR) in uncontrolled environment targets to recognize different actions from a video. Effective HAR model can be employed to application like human-computer interaction, health care, person tracking and video surveillance. Machine Learning (ML) approaches, specifically, Convolutional Neural Network (CNN) models had been widely used and achieved impressive results through feature fusion. Accuracy and effectiveness of these models continue to be the biggest challenge in this field. In this article, a novel feature optimization algorithm, called improved Shark Smell Optimization (iSSO) is proposed to reduce the redundancy of extracted features. This proposed technique is inspired from the behavior of white sharks, how they find the best prey in the whole search space. The proposed iSSO algorithm divides the Feature Vector (FV) into subparts, where a search is conducted to find local optimal features from each subpart of FV. Once local optimal features are selected, a global search is conducted to further optimize these features. The proposed iSSO algorithm is employed on nine (9) selected CNN models. These CNN models are selected based on their top-1 and top-5 accuracy in ImageNet competition. To evaluate the model, two publicly available datasets UCF-Sports and Hollywood2 are selected.**

## I. INTRODUCTION

Human Action Recognition (HAR) includes the action recognition of a person through imaging data which has various applications. Recognition approaches can be divided into three categories: multi-model, overlapping categories, and video sequences [1]. This data used for recognition is the major difference among images and video categories. Data in form of images and videos are acquired through cameras in controlled and uncontrolled environments. With the advancement of technology in past decades, various smart devices have been developed which to collect images and video data for HAR, health monitoring and disease prevention [2]. Different research carried out on HAR through images or videos over the last three decades [3, 4]. Human visual systems get visual information about an object such as its movement, shape, and its variations. This information is used to investigate the biophysical processes of HAR. Computer vision systems have achieved a very good accuracy while catering to different challenges such as occlusion, background clutter, scale and rotation invariance, and environmental changes [5].

HAR depending upon the action complexity can be divided into primitive, single-person, interaction, and group action recognition [6]. The basic movement of a single human body part considers in primitive action, a set of primitive actions of one person includes in single-person action, a collection of humans and objects involves in interaction while collective actions performed by a group of people are group actions. Computer vision based HAR systems are divided into hand crafted feature-based methods and deep learning-based methods. The combined framework of hand-crafted and deep features are also employed by many researchers [7].

The data plays an important role in efficient HAR systems. The HAR data is categorized into color channels, depth, and skeleton information. Texture information can be extracted from color channels i.e., RGB which is closed to the visual appearance, but illumination variations can affect the visual data [8]. Depth map information is invariant to the lighting changes which is helpful in foreground object extractions. 3D information can also be captured through a depth map, but the noise factors should be considered while capturing the depth map. Skeletons information can be gathered through color channels and depth maps, but it can be exploited from the environmental factors [9]. HAR systems use different levels of features such as whole data as the input of HAR used in [10]. Apart from features, motion is an important factor that can be incorporated into the feature computation step. It includes optical flow for capturing low-level feature information in multiple video frames. Some researchers included motion information in the classification step with Conditional Random Fields, Hidden Markov Models, Long-Short Term Memory (LSTM), Recurrent Neural Network (RNN), and 3D Convolutional Neural Network (CNN) [11-15]. These HAR systems have good recognition accuracy using the most appropriate feature set.

A CNN-based convolutional 3D (C3D) network was proposed in [16]. The major difference between the 3D CNN and the proposed was that it utilized the whole video as an input

instead of a few frames or segmented frames, which makes it robust for large databases. The architecture of the C3D network comprises several layer groups like convolutional layers = 8, maximum pooling layers = 5, fully connected layers = 2, and the last softmax loss layer. UCF 101 dataset was utilized to evaluate the best combination of proposed network architecture. The best performance achieved by the proposed network was using a 3×3×3 convolutional filter without updating the other parameter. The researcher came up with RNNs [17] to overcome the limitation action of CNN models of information derivation from long timelapse. RNN has proved robust while extracting time dimension features and has one drawback of gradient disappearance. The mentioned problem is addressed by presenting Long Short-Term Memory Network (LSTM) [18], which utilizes processors to gauge the information integrity and relevance. Normally, input gate, output gate, and forget gates are utilized in the processor. The information flow is controlled by gates in the processor and unnecessary information which required large memory chunks is stored for long-term tasks.

A ConvNet architecture for spatiotemporal fusion of video fragments has evaluated its performance on dataset UCF-101 by achieving accuracy of 93.5% and HMDB-51 by achieving accuracy of 69.2% [19]. An architecture is proposed to handle 3D signals effectively and efficiently and introduced Factorized Spatio-Temporal Convolutional Network (FSTCN). It was tested on two publicly available dataset UCF-101 and achieved 88.1% accuracy, while achieved 59.0% accuracy on HMDB-51 [20]. In another method, LSTM models are trained to utilize the differential gating scheme, which focuses on the varying gain due to the slow movements between the successive frames, change based on Derivate of States (DoS) and this combined called as differential RNN (dRNN). Method is implemented on KTH and MSRAction3D datasets. The accuracy achieved on their datasets are 93.96% and 92.03% respectively [21].

This article presents an improved form of Shar Smell Algorithm (SSO), which reduces the redundant features. Proposed algorithm utilizes both, SSO and White Shark Optimization (WSO) properties to solve the redundancy issues. Proposed iSSO divides the population into sub-spaces to find local and global optimal features. At the end, these extracted local features are used to optimize global features. Features are extracted using 9 pre-trained CNN models, which are selected based on their top-1 and top-5 accuracies in ImageNet competition. This model is tested on two publicly available datasets UCF-Sports (D1) and Hollywood2 (D2) and it have obtained better results than state-of-the-art (SOTA) methods.

## II. PROPOSED METHODOLOGY

In uncontrolled environment, various viewports, illuminations and changing backgrounds, traditional hand-crafted features have been proved insufficient [22]. In the age of big data and evolution of ML methods, Deep Learning (DL) has achieved remarkable results[23-25]. These results have motivated the researchers around the globe to apply these DL methods on domains involving video data. After the 2012, challenge of ImageNet classification has drastically changed the dimensions of DL methods, when CNNs made a huge breakthrough. The main difference between CNN methods and local feature-based methods is that CNN iteratively and automatically extracts deep features through its interconnected layers.

### *A. Transfer Learning of Pre-Trained CNN Models*

Artificial Intelligence (AI) and Machine Learning (ML) has a sub-domain, called Transfer Learning (TL), which transforms the learned knowledge of one problem (base problem) to another problem (target problem). TL basically improves the learning of a model through the data provided for the target problem. A model trained to classify Wikipedia text can be utilized to classify the texts of simple documents after TL. A model trained to classify cards can also classify birds. The nature of these problem is same, which is to classify objects. TL provides scalability to a trained model, which enables it to recognize different type of objects. Since 2015, after the first CNN model, AlexNet [22] was proposed, a lot of CNN architectures were proposed. The base for all these models was a competition, where a dataset, ImageNet [ref], having 1000 classes was presented. The efficiency of all proposed CNN models till date, is still measured on how the proposed model performs on ImageNet dataset. In this research, nine of the most used CNN models are selected, where, through TL, features of input images from selected datasets will be extracted. Tab. 1 lists all selected CNN models along with their depth, size, input size, number of parameters and their top-1 and top-5 accuracies on ImageNet datasets.

TABLE I DIFFERENT CHARACTERISTICS OF SELECTED PRE-TRAINED CNN MODELS

| Model | Depth | Size (MB) | Input Size | Parameters (Millions) | Accuracy (%) | |
|---|---|---|---|---|---|---|
| | | | | | Top-1 | Top-5 |
| VGG19 (Vg) [23] | 19 | 535 | 224×224×3 | 143.6 | 71.3 | 90.0 |
| MobileNetV2 (Mo) [24] | 53 | 13 | 299×299×3 | 3.5 | 71.3 | 90.1 |
| Resnet50 (Re) [25] | 50 | 96 | 224×224×3 | 25.6 | 74.9 | 92.1 |
| EfficientNet-B0 (Ef) [26] | 82 | 20 | 224×224×3 | 5.3 | 77.1 | 93.3 |
| DarkNet53 (Da) [27] | 53 | 155 | 256×256×3 | 41.6 | 77.2 | 93.8 |
| DenseNet201 (De) [28] | 201 | 77 | 224×224×3 | 20.0 | 77.3 | 93.6 |
| Xception (Xe) [29] | 71 | 85 | 299×299×3 | 22.9 | 79.0 | 94.5 |
| InceptionResNetV2 (In) [30] | 164 | 209 | 299×299×3 | 55.9 | 80.3 | 95.3 |
| NASNetLarge (Na) [31] | - | 332 | 331×331×3 | 88.9 | 82.5 | 96.0 |

The structure of all these selected pre-trained models is different because of the nature and arrangement of layers. Selected feature extraction layer and extracted features per image vary from model-to-model. For Vg, fc7 layer is selected to extract 4096 features for a single image. 1280 and 4032 features are extracted from the global_average_pooling2d_1 and global_average_pooling2d_2 layers of Mo and Na models, respectively. avg_pool is selected as a feature extraction layer for Re, De, Xe and In models, which extracted 2048, 1920, 2048 and 1536 features, respectively. avg1 is selected as feature extraction layer for Da, and it extracted 1024 features against a

single image. When Ef model is used as feature extractor, it extracts 1280 features from the GlobAvgPool layer. All these extracted features are forwarded to iSSO for optimization.

### *B. Improved Shark Smell Optimization (iSSO)*

The meta-heuristic model used in this article is an improved form of Shark Smell Optimization (SSO) [32]. The SSO was proposed after inspiration was taken from the species of sharks. Sharks are considered as most hazardous and strongest predacious of the universe [33]. Sharks are creatures having keen ability of smell, highly contrasted vision due to the sturdy eyesight and powerful muscles. They have more than 300 sharp, pointing, and triangular teeth in their gigantic jaws. Sharks usually strike with a large and abrupt bite on prey, which proves so sudden that the prey cannot avoid it. These sharks hunt the prey by using their extreme sense of smelling and hearing the traits of prey. The iSSO algorithm initially divides the whole search space in ş subparts. The algorithm, then performs the local and global search to find the optimum prey in both, local and global search spaces of ş. Once an optimum prey is located, the search then continues to find all the optimal preys in remaining subparts. The process mentioned below is for a single subpart. The whole process will be repeated for all ş. Other factor is quantity of selected optimal features. For this, Ŧ denotes the total selected features.

### *C. Prey Tracking*

Sharks wander in the ocean freely just like any other organism of the sea and search for the prey. In that search, sharks update their positions in accordance with the traits of prey. They apply all their tricks to locate, stalk and track down the prey. All senses of sharks along with their average distance range are illustrated in Fig. 1. All these illustrated features help them to exploit and search the whole space for hunting the prey.

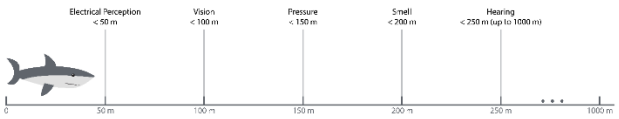



Fig. 1. Senses of shark along with its average distance range

## III. EXPERIMENTAL RESULTS

The proposed iSSO algorithm is evaluated by performing multiple experiments under different parameters, which efficiently verifies the performance of this algorithm. This section provides an in-depth view of performed experiments along with ablation analysis and comparison with existing techniques.

### *D. Experimental Setup and Datasets*

The proposed iSSO algorithm is evaluated on two (2) benchmark datasets including UCF-Sports Dataset (D1) [34] and Hollywood2 Dataset (D2) [35]. D1 contains a total of 150 videos from 10 classes are included in this dataset, which represents human actions in different viewpoints and range of scenes. D2 contains a total of 1,707 videos across 12 classes. These videos are extracted from 69 Hollywood movies.

The proposed iSSO model is trained, tested, and validated using HP Z440 workstation having NVIDIA Quadro K2000 with a GPU memory of 2GB DDR5. This card has 382 CUDA cores along with 128-bit memory interface and 173 GB/s memory bandwidth. MATLAB2021a was used for training, testing, and validation. All selected pre-trained models are transfer learned with initial learning rate of 0.0001 with an average decrease of 5% after 7 epochs. The whole process has 160 epochs and an overall momentum of 0.45. Selected datasets are split using the standard 70-15-15 ratio for training, testing, and validation. During the testing of proposed model, eight (8) classifiers were trained, which include Bagged Tree (BTree), Linear Discriminant Analysis (LDA), three kernels of k-Nearest Neighbor (kNN), i.e., Ensemble Subspace kNN (ES-kNN), Weighted kNN (W-kNN) and Fine kNN (F-kNN), and three kernels of Support Vector Machine (SVM) i.e., Cubic SVM (C-SVM), Quadratic SMV (Q-SVM) and Multi-class SVM (M-SVM). The performance of proposed iSSO algorithm is evaluated using six metrices, such as Sensitivity (Sen), Correct Recognition Rate (CRR), Precision (Pre), Accuracy (Acc), Prediction Time (PT) and Training Time (TT). All experimental results presented in next section are achieved after performing each experiment at least five times, using the same environment and factors.

### *E. Recognition Results*

The efficiency of proposed model is evaluated by performing multiple experiments. Initially, impact of all selected pre-trained models is noted by feeding the dataset and extracting features from the selected output layer. In the next experiment, proposed iSSO algorithm is employed on extracted deep features. And finally, the iSSO enabled CNN model with highest accuracy is further forwarded to the other classifiers. It is noteworthy that all the selected classifiers were used during this experiment, but F-kNN achieved the highest accuracy, thus Tab. 2 contains the results of F-kNN. While using D1, Na model achieved highest average Acc of 97.44 was achieved. This average accuracy has a factor, ±1.36%, by which it alters during the five experiments. Similarly, Na obtained 96.97% CRR. The F-kNN took 206 minutes on average to train and 0.53 seconds to predict an input image. The lowest average Acc of 73.02% was obtained by Vg model, whereas Ef took the highest TT of 347 minutes.

TABLE II PERFORMANCE OF ISSO ON SELECTED CNN MODELS ON D1

| CNN Model | iSSO | | Acc (%) | CRR (%) | TT (m) | PT (s) |
|---|---|---|---|---|---|---|
| | No | Yes | | | | |
| Vg | ✓ | | 73.02 ± 1.56 | 73.76 ± 2.06 | 240 ± 18 | 0.76 ± 0.29 |
| | | ✓ | 75.53 ± 2.96 | 75.95 ± 1.27 | 209 ± 21 | 0.59 ± 0.15 |
| Mo | ✓ | | 79.18 ± 4.59 | 78.78 ± 1.74 | 312 ± 36 | 0.72 ± 0.32 |
| | | ✓ | 82.95 ± 1.59 | 83.28 ± 2.06 | 277 ± 22 | 0.57 ± 0.12 |
| Re | ✓ | | 83.28 ± 2.73 | 83.58 ± 2.02 | 267 ± 16 | 0.78 ± 0.23 |
| | | ✓ | 86.71 ± 1.72 | 86.76 ± 2.11 | 266 ± 26 | 0.59 ± 0.29 |
| Ef | ✓ | | 74.51 ± 2.07 | 74.81 ± 2.56 | 347 ± 28 | 0.82 ± 0.19 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | ✓ | 77.68 ± 1.74 | 76.16 ± 1.33 | 300 ± 19 | 0.55 ± 0.25 |
| Da | ✓ | | 88.86 ± 2.82 | 89.25 ± 1.77 | 319 ± 21 | 0.79 ± 0.22 |
| | | ✓ | 91.84 ± 4.61 | 90.35 ± 1.28 | 265 ± 12 | 0.55 ± 0.39 |
| De | ✓ | | 92.56 ± 1.64 | 92.98 ± 1.75 | 267 ± 22 | 1.15 ± 0.22 |
| | | ✓ | 94.74 ± 1.32 | 95.04 ± 1.87 | 234 ± 33 | 0.72 ± 0.28 |
| Xe | ✓ | | 82.22 ± 2.18 | 82.68 ± 1.68 | 258 ± 26 | 0.87 ± 0.33 |
| | | ✓ | 84.65 ± 2.20 | 84.96 ± 2.48 | 230 ± 40 | 0.66 ± 0.24 |
| In | ✓ | | 80.83 ± 1.84 | 81.27 ± 1.82 | 225 ± 30 | 1.04 ± 0.38 |
| | | ✓ | 82.26 ± 2.09 | 83.89 ± 1.45 | **194 ± 10** | 0.76 ± 0.33 |
| **Na** | ✓ | | 95.42 ± 1.59 | 94.94 ± 2.42 | 254 ± 28 | 0.71 ± 0.31 |
| | | ✓ | **97.44 ± 1.36** | **96.97 ± 1.82** | 206 ± 30 | **0.53 ± 0.24** |

Once a model with best performance is selected in the first experiment, this model is used to train all selected classifiers. As mentioned earlier. F-kNN performed better on D1 when Na is selected as the base CNN model. This classifier achieved average Sen of 97.37%, average CRR of 96.97% and Pre of 97.28%. The second best average Acc of 91.75% was achieved by Es-kNN. The worst performing classifier was BTree, which could only achieve 80.83% average Acc. The lowest average TT was of 193 seconds and lowest average PT of 0.39 seconds was taken by LDA, but it could only achieve 84.16% Acc.

The proposed model is also evaluated on D2, where Da network achieved maximum average Acc of 80.66%. The change factor of this model is 1.04%, after performing the same experiment 5 times. The average CRR of this model is noted at 79.68%. The best classifier for this model is M-SVM, which took 139 minutes on average to train and 0.48 seconds on average to predict an input image. Second best average Acc of 78.27% is achieved by De, which also achieves 78.66% CRR. For this model, M-SVM took 221 minutes to train and 0.54 seconds to predict. The lowest average accuracy of 60.02% on D2 is again achieved by Vg, where the selected classifier took 297 minutes to train and 1.45 seconds to predict an input image. The performances of all selected CNN models with and without iSSO algorithm are compared in Tab. 3.

TABLE III. PERFORMANCE OF ISSO ON SELECTED CNN MODELS ON D2

| **CNN Model** | **iSSO** | | **Acc (%)** | **CRR (%)** | **TT (m)** | **PT (s)** |
|---|---|---|---|---|---|---|
| | **No** | **Yes** | | | | |
| Vg | ✓ | | 60.02 ± 1.64 | 60.35 ± 2.63 | 297 ± 25 | 1.45 ± 0.17 |
| | | ✓ | 63.25 ± 1.07 | 62.96 ± 2.08 | 224 ± 17 | 0.82 ± 0.14 |
| Mo | ✓ | | 74.14 ± 1.65 | 73.85 ± 2.13 | 292 ± 46 | 0.88 ± 0.31 |
| | | ✓ | 76.83 ± 2.17 | 76.23 ± 2.38 | 236 ± 28 | 0.55 ± 0.12 |
| Re | ✓ | | 71.61 ± 2.35 | 70.19 ± 2.31 | 245 ± 28 | 1.19 ± 0.15 |
| | | ✓ | 73.82 ± 2.19 | 74.86 ± 1.96 | 176 ± 32 | 0.88 ± 0.26 |
| Ef | ✓ | | 68.45 ± 2.16 | 67.54 ± 2.33 | 230 ± 12 | 0.99 ± 0.28 |
| | | ✓ | 72.12 ± 1.54 | 70.69 ± 2.82 | 192 ± 21 | 0.63 ± 0.34 |
| Da | ✓ | | 77.22 ± 1.36 | 78.69 ± 2.09 | 161 ± 42 | 0.65 ± 0.24 |
| | | ✓ | **80.66 ± 1.04** | **79.68 ± 1.21** | **139 ± 32** | **0.48 ± 0.22** |
| De | ✓ | | 75.23 ± 2.62 | 75.22 ± 2.76 | 254 ± 44 | 0.79 ± 0.39 |
| | | ✓ | 78.27 ± 1.93 | 78.66 ± 1.15 | 221 ± 27 | 0.54 ± 0.37 |
| Xe | ✓ | | 63.04 ± 1.95 | 63.33 ± 2.17 | 212 ± 24 | 0.73 ± 0.21 |
| | | ✓ | 66.86 ± 2.27 | 66.38 ± 2.68 | 180 ± 27 | 0.53 ± 0.14 |
| In | ✓ | | 66.14 ± 2.12 | 67.14 ± 2.39 | 236 ± 14 | 0.74 ± 0.16 |
| | | ✓ | 69.24 ± 2.04 | 68.23 ± 2.29 | 204 ± 24 | 0.56 ± 0.24 |
| **Na** | ✓ | | 71.28 ± 2.89 | 71.31 ± 2.67 | 364 ± 26 | 0.61 ± 0.23 |
| | | ✓ | 75.62 ± 2.38 | 74.76 ± 1.86 | 292 ± 48 | 0.46 ± 0.15 |

After the selection of best performing CNN model, all selected classifiers are trained on the extracted features of that CNN model. During this experiment, selected evaluation matrices are used to note the performance of each classifier. M-SVM has achieved the best average Sen of 79.22%, best average CRR of 79.68%, best Pre of 79.84% and best average Acc of 80.66%. This classifier requires 280 minutes for training and 0.48 seconds for predicting an input image. The second best average Acc of 75.88% is obtained by W-kNN, which took 280 minutes to train and 0.36 seconds to predict. The lowest TT is noted at 115 minutes for BTree, but the achieved average Acc is 50.95%.

## IV. CONCLUSIONS

In this article, an analysis of pre-trained CNN models is presented, where 9 models are selected based on their total parameters, size and Top-1 and Top-5 accuracies. These selected pre-trained CNN models are trained on the selected dataset using the TL. The output layer of these pre-trained models is mentioned, and no experiments are performed based on selection of output layer. The extracted features of these CNN models are forwarded to proposed iSSO, which is an improved algorithm from the traditional SSO. The iSSO algorithm divides the feature vector into subsets, where each subset is then used to find the local and global best features. The selection of local and global best features is inspired by the searching capabilities of white shark, which uses its senses to find the optimal prey. Once the features are selected, the results are taken using selected publicly available datasets. The limitation of this work is the training time, which is too high, i.e., lowest training time for D1 is 194 minutes and for D2, it is 139 minutes. The one reason of taking this much TT is the dataset, which includes videos. But the main reason is the architecture of these models, which have too many repeated blocks of layers, which can be reduced. In future, the

architecture of best performing CNN models of this article will be analyzed to detect and reduce the repeated blocks of layers. The impact of these repeated blocks can also be analyzed.

## ACKNOWLEDGMENT

This This work was supported by the Collabo R&D between Industry, Academy, and Research Institute(S3250534) funded by the Ministry of SMEs and Startups (MSS, Korea), and the Soonchunhyang University Research Fund.

## REFERENCES

[1] V. Sharma, M. Gupta, A. K. Pandey, D. Mishra and A. Kumar, "A review of deep learning-based human activity recognition on benchmark video datasets," Applied Artificial Intelligence, vol. 36, no.1, pp. 1-17, 2022.

[2] S. K. Yadav, K. Tiwari, H. M. Pandey and S. A. Akbar, "A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions," Knowledge-Based Systems, vol. 223, no.11, pp. 51-83, 2021.

[3] I. M. Nasir, M. Raza, J. H. Shah, S.-H. Wang, U. Tariq et al,, "HAREDNet: A deep learning based architecture for autonomous video surveillance by recognizing human actions," Computers and Electrical Engineering, vol. 99, no.1, pp. 1-16, 2022.

[4] M. Raza, J. H. Shah, M. A. Khan and A. Rehman, "Human action recognition using machine learning in uncontrolled environment," Artificial Intelligence and Data Analytics, vol. 12, no.3, pp. 182-187, 2021.

[5] P. Pareek and A. Thakkar, "A survey on video-based human action recognition: recent updates, datasets, challenges, and applications," Artificial Intelligence Review, vol. 54, no.3, pp. 2259-2322, 2021.

[6] A. Sarkar, A. Banerjee, P. K. Singh and R. Sarkar, "3D human action recognition: through the eyes of researchers," Expert Systems with Applications, vol. 1, no.2, pp. 11-42, 2022.

[7] I. U. Khan, S. Afzal and J. W. Lee, "Human activity recognition via hybrid deep learning based model," Sensors, vol. 22, no.1, pp. 323-349, 2022.

[8] Z. Fu, X. He, E. Wang, J. Huo, J. Huang et al., "Personalized human activity recognition based on integrated wearable sensor and transfer learning," Sensors, vol. 21, no.3, pp. 885-903, 2021.

[9] H. Wang, B. Yu, K. Xia, J. Li and X. Zuo, "Skeleton edge motion networks for human action recognition," Neurocomputing, vol. 423, no.3, pp. 1-12, 2021.

[10] [10]Y. B. Cheng, X. Chen, D. Zhang and L. Lin, "Motion-transformer: self-supervised pre-training for skeleton-based action recognition," Multimedia Tools and Applications, vol. 4, no.1, pp. 1-32, 2021.

[11] K. Liu, L. Gao, N. M. Khan, L. Qi and L. Guan, "Integrating vertex and edge features with graph convolutional networks for skeleton-based action recognition," Neurocomputing, vol. 466, no.13, pp. 190-201, 2021.

[12] T. Xue and H. Liu, "Hidden markov model and its application in human activity recognition and fall detection: A review," Communications, Signal Processing, and Systems, vol. 1, no.1, pp. 863-89, 2022.

[13] C. Yin, J. Chen, X. Miao, H. Jiang and D. Chen, "Device-free human activity recognition with low-resolution infrared array sensor using long short-term memory neural network," Sensors, vol. 21, no.10, pp. 3551-3561, 2021.

[14] A. Anagnostis, L. Benos, D. Tsaopoulos, A. Tagarakis, N. Tsolakis et al., "Human activity recognition through recurrent neural networks for human–robot interaction in agriculture," Applied Sciences, vol. 11, no.5, pp. 2188-2205, 2021.

[15] W. Ding, C. Ding, G. Li and K. Liu, "Skeleton-based square grid for human action recognition with 3D convolutional neural network," IEEE Access, vol. 9, no.3, pp. 54078-54089, 2021.

[16] D. Tran, L. Bourdev, R. Fergus, L. Torresani and M. Paluri, "Learning spatiotemporal features with 3d convolutional networks," Image and Vision Computing, vol. 13, no.5, pp. 4489-4497, 2019.

[17] A. Graves, A. R. Mohamed and G. Hinton, "Speech recognition with deep recurrent neural networks," Acoustics, Speech and Signal Processing, vol. 1, no.1, pp. 6645-6649, 2013.

[18] A. Graves, "Long short-term memory," Supervised Sequence Labelling with Recurrent Neural Networks, 1st ed., vol. 385. Berlin, DEU: Springer Berlin, pp. 37–45, 2012.

[19] C. Feichtenhofer, A. Pinz and A. Zisserman, "Convolutional two-stream network fusion for video action recognition," Computer Vision and Pattern Recognition, vol. 2, no.1, pp. 1933-1941, 2012.

[20] L. Sun, K. Jia, D. Y. Yeung and B. E. Shi, "Human action recognition using factorized spatio-temporal convolutional networks," Computer Vision and Applications, vol. 1, no.1, pp. 4597-4605, 2015.

[21] V. Veeriah, N. Zhuang and G. J. Qi, "Differential recurrent neural networks for action recognition," Computer Vision amd Applications, vol. 1, no.1, pp. 4041-4049, 2015.

[22] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh et al., "Imagenet large scale visual recognition challenge," International Journal of Computer Vision, vol. 115, no.3, pp. 211-252, 2015.

[23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," International Journal of Computer Vision, vol. 115, no.3, pp. 1409.1556, 2014.

[24] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. C. Chen, "Mobilenetv2: inverted residuals and linear bottlenecks," Computer Vision and Pattern Recognition, vol. 2, no.1, pp. 4510-4520, 2018.

[25] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," Computer Vision and Pattern Recognition, 2016, vol. 1, no.13, pp. 770-778, 2016.

[26] M. Tan and Q. Le, "Efficientnet: rethinking model scaling for convolutional neural networks," Machine Learning Tools, vol. 1, no.1, pp. 6105-6114, 2019.

[27] J. Redmon, "Darknet: open source neural networks in c," 1st ed., vol. 3. New York, USA: Springer, pp. 152–183, 2013.

[28] G. Huang, Z. Liu, L. V. D. Maaten and K. Q. Weinberger, "Densely connected convolutional networks," Computer Vision and Pattern Recognition, vol. 3, no.12, pp. 4700-4708, 2017.

[29] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in Computer Vision and Pattern Recognition, vol. 3, no.11, pp. 1251-1258, 2017.

[30] C. Szegedy, S. Ioffe, V. Vanhoucke and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in Artificial Intelligence, vol. 1, no.1, pp. 1-13, 2017.

[31] B. Zoph, V. Vasudevan, J. Shlens and Q. V. Le, "Learning transferable architectures for scalable image recognition," Computer Vision and Pattern Recognition, vol. 3, no.11, pp. 8697-8710, 2018.

[32] O. Abedinia, N. Amjady and A. Ghasemi, "A new metaheuristic algorithm based on shark smell optimization," Complexity, vol. 21, no.5, pp. 97-116, 2016.

[33] S. Wroe, D. R. Huber, M. Lowry, C. McHenry, K. Moreno et al., "Three-dimensional computer analysis of white shark jaw mechanics: how hard can a great white bite?," Journal of Zoology, vol. 276, no.4, pp. 336-342, 2008.

[34] K. Soomro and A. R. Zamir, "Action recognition in realistic sports videos," in Computer Vision in Sports: Springer, vol. 1, no.1, pp. 181-208, 2014.

[35] M. Marszalek, I. Laptev and C. Schmid, "Actions in context," Computer Vision and Pattern Recognition, vol. 1, no.1, pp. 2929-2936, 2009.

[36] J. Zhao, "Sports motion feature extraction and recognition based on a modified histogram of oriented gradients with speeded up robust features," Journal of Computers, vol. 33, no.1, pp. 63-70, 2022.

[37] A. A. Alibari, J. S. Alzahrani, A. Qahmash, M. Maray, M. Alghamdi et al., "Quantum water strider algorithm with hybrid-deep-learning-based activity recognition for human–computer interaction," Applied Sciences, vol. 12, no.14, p. 6848, 2022.

[38] K. Muhammad, A. Ullah, A. S. Imran, M. Sajjad, M. Kiran et al., "Human action recognition using attention based LSTM network with dilated CNN features," Future Generation Computer Systems, vol. 125, no.1, pp. 820-830, 2021.

[39] S. Kiran, M. A. Khan, M. Y. Javed, M. Alhaisoni, U. Tariq et al., "Multi-layered deep learning features fusion for human action recognition," Computers, Materials & Continua, vol. 69, no.3, pp. 4061-4075, 2021.

[40] F. Afza, M. A. Khan, M. Sharif, S. Kadry, G. Manogaran et al., "A framework of human action recognition using length control features fusion and weighted entropy-variances based feature selection," Image and Vision Computing, vol. 106, no.1, pp. 104090-104105, 2021.

[41] W. Zou, S. Zhuo, Y. Tang, S. Tian, X. Li et al., "STA3D: spatiotemporally attentive 3D network for video saliency prediction," Pattern Recognition Letters, vol. 147, no.3, pp. 78-84, 2021.

[42] J. Chen, Z. Li, Y. Jin, D. Ren and H. Ling, "Video saliency prediction via spatio-temporal reasoning," Neurocomputing, vol. 462, no, 2, pp. 59-68, 2021.

[43] C. A. Aly, F. S. Abas and G. H. Ann, "Robust video content analysis schemes for human action recognition," Science Progress, vol. 104, no.2, pp. 5480-5501, 2021.

[44] A. Ullah, K. Muhammad, W. Ding, V. Palade, I. U. Haq et al., "Efficient activity recognition using lightweight CNN and DS-GRU network for surveillance applications," Applied Soft Computing, vol. 103, no. 1, pp. 107-122, 2021.

[45] C. Wu, Y. Li, Y. Zhang and B. Liu, "Double constrained bag of words for human action recognition," Signal Processing: Image Communication, vol. 98, no. 3, pp. 1163-1199, 2021.

# Design of Evolutionary Algorithm based Unequal Clustering for Energy Aware Wireless Sensor Networks

Mohammed Altaf Ahmed[1], T. Satyanarayana Murthy[2], Fayadh Alenezi[3], E. Laxmi Lydia[4], Seifedine Kadry[5], Yena Kim[6] and Yunyoung Nam[6, *]

[1]*Department of Computer Engineering, College of Computer Engineering & Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia*
[2]*Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India.*
[3]*Department of Electrical Engineering, College of Engineering, Jouf University, Saudi Arabia.*
[4]*Department of Computer Science and Engineering, Vignan's Institute of Information Technology, Visakhapatnam, 530049, India.*
[5]*Department of Applied Data Science, Noroff University College, Kristiansand, Norway*
[6]*Department of ICT Convergence, Soonchunhyang University, South Korea*

**Contact: Yunyoung Nam (ynam@sch.ac.kr)*

*Abstract*— **Wireless Sensor Networks (WSN) plays a vital role in several real time applications ranging from military to civilian applications. Despite the benefits of WSN, energy efficiency becomes a major part of challenging issue in WSN, which necessitate proper load balancing amongst the clusters and serves a wider monitoring region. Clustering technique for WSN has several benefits namely lower delay, higher energy efficiency, and collision avoidance. But clustering protocol has number of challenges. In a largescale of network, cluster-based protocols mainly adapt multi-hop routing for saving energy that leads to hot spot problems. A hot spot problem becomes a problem where cluster node nearer to the base station (BS) tends to drain the energy much quicker than other nodes because of the need to implement more transmission. This article introduces a Jumping Spider Optimization Based Unequal Clustering Protocol for Mitigating Hotspot Problem (JSOUCP-MHP) in WSN. The JSO algorithm is stimulated by the characteristics of spiders naturally and mathematically modelled the hunting mechanism such as search, persecution, and jumping skills to attack prey. The presented JSOUCP-MHP technique mainly resolves the hot spot issue for maximizing the network lifespan. To attain this, the JSOUCP-MHP technique elects proper set of cluster heads (CHs) using average residual energy (RE). In addition, the JSOUCP-MHP technique determines the cluster sizes based on two measures i.e., RE and distance to BS (DBS). The proposed JSOUCP-MHP technique is examined under several experiments to ensure its supremacy. The comparison study shows the significance of the JSOUCP-MHP technique over other models.**

## I. INTRODUCTION

Wireless Sensor Networks (WSNs) are advanced for monitoring and sensing vital signs of area and environment by making use of linked and distributed sensor nodes (SNs). The SNs were divided into Gateway Nodes (GN), normal, and sink nodes [1]. The SNs were small in size and have insufficient sources with regard to processing, storage, and space energy [2]. In many cases, the SNs were deployed in a very harsh and intense ecosystem. The first and foremost goal of this positioning was sense information distantly and sends it to system or user for decision making. In order to send data, there comes a demand for more effective systems for managing the energy system of nodes and enhancing network lifetime [3]. The SNs perceiving and sensing the data from nearby environment and making processes and transmit it to the nearby nodes until data arrived at the base station (BS). In WSN, because of the restricted energy sources of SNs, there comes a crucial need for well-effective and balanced data aggregation system and energy efficient routing protocol [4]. The energy feature becomes one such factor to be considered for devising any solution for WSNs. Numerous types of routing protocols were devised for conserving the energy of SNs [5]. The clustering approach becomes an effective topological control system that could efficiently enhance the scalability duration and lifetime of WSNs. The WSN applications have achieved popularity including health monitoring, target tracking, disaster response, environmental monitoring, and security [6].

In a clustering ecosystem, data can be involved from one node to another, and occurs energy holes or hotspot difficulties [7]. A hotspot can be made by SNs positioned nearby the BS and rapidly drain energy because traffic arises from other nodes and is sent from it. Such SNs not just transmit their data but also sends the data from other bases because the initial death of nodes causes hotspot problems. The methods namely unequal clustering methods and mobile sinks or mobile data mules were majorly utilized for resolving this type of problem [8]. But many unequal clustering techniques devised for the solution of the hotspot leverages cluster which smaller in size adjacent to BS and cluster size rises case when we were going far away from BS. The cluster size was inversely proportional to the distance from BS. The clusters adjacent to BS hold a greater quantity of nodes that is helpful in efficient load sharing [9]. The projected method for size allotment and size difference of clusters leads to reduction of frequency where a specific node

turns out to be cluster head (CH) [10]. This aids in maintenance of the overall connectivity and thwarts network isolation. In line with this, hot spot issues can be reduced. One of the problem with regard to unequal clustering was cluster size decreasing or increasing ratio that was not evidently conferred in prevailing methods.

Arikumar et al. [11] modelled an Energy Efficient LifeTime Maximization (EELTM) technique that uses intelligent techniques Fuzzy Inference System (FIS) and particle swarm optimization (PSO). In addition, devised an optimal CH–CR selecting method in this technique that exploits fitness values (FV) computed by the PSO approach for determining 2 optimally nodes in every cluster to serve as Cluster Router (CR) and CH. The CH which is selected exclusively accumulates the data from its cluster members, although the CR is accountable to receive the data which is collected from its CH and sending it to BS. Mehra [12] formulated an enhanced fuzzy unequal clustering and routing protocol (E-FUCA) in which vital variables were regarded at the time of CH candidate selecting, and intellectual decision utilizing FL was considered by non-CH nodes at the time of election of its CH for forming clusters. To expand the lifespan, FL is utilized as next-hop choice for effective routing. And carried out the simulation experimentations for 4 cases and comparison is made with the propound performance of protocol with recent similar protocols.

Wang and Hu [13] introduced an energy-efficient unequal clustering routing algorithm (UCRA). At first, the monitoring region has been separated by concentric circles as rings of various sizes. Then, the CH was selected relating to RE and position. Before clustering, every typical sensor links cluster related to the electability of every CH. Nguyen et al. [14] presented a new optimization technique, like the compact bat algorithm (cBA), for using it for class of optimization issues which involves gadgets containing inadequate hardware sources. A real-valued prototype vector can be employed for probabilistic functions for generating every candidate solution for optimizing cBA. The projected cBA was widely assessed on numerous continual multi-modal functions and the unequal clustering of WSN (uWSN) issues.

In [15], competitive swarm optimization (CSO) related methods were introduced, together called such methods as CSO-UCRA. Initially, the CH selecting approach was provided relies on CSO related approach, after assigning non-CH sensors to CHs. At last, a CSO related routing method was provided. New fitness functions and efficient particle encoding techniques were advanced for such methods. Guleria and Verma [16] devise the new ant colony meta-heuristic related unequal clustering for novel CH election. The neighbor finding level and link maintenance by Meta-Heuristic Ant Colony Optimization (ACO) technique choose the optimal path among the nodes that rises the packets distributed to the destiny.

This article introduces a Jumping Spider Optimization Based Unequal Clustering Protocol for Mitigating Hotspot Problem (JSOUCP-MHP) in WSN. The JSO algorithm is stimulated by the characteristics of spiders in nature and mathematically modelled the hunting mechanism such as search, jumping skills, and persecution to attack prey. The presented JSOUCP-MHP technique mainly resolves the hot spot issue for maximizing the network lifespan. To attain this, the JSOUCP-MHP technique elects proper set of cluster heads (CHs) using average RE. In addition, the JSOUCP-MHP technique determines the cluster sizes based on two measures i.e., RE and distance to BS (DBS). The proposed JSOUCP-MHP technique is examined under several experiments to ensure its supremacy.

## II. The Proposed Unequal Clustering Protocol

In this article, a novel JSOUCP-MHP algorithm was presented for resolving hot spot issues in WSN. The JSO algorithm is stimulated by the characteristics of spiders in nature and mathematically modelled the hunting mechanism such as search, jumping skills, and persecution to attack prey. The presented JSOUCP-MHP technique mainly resolves the hot spot issue for maximizing network lifespan. Fig. 1 depicts the overall process of JSOUCP-MHP approach.

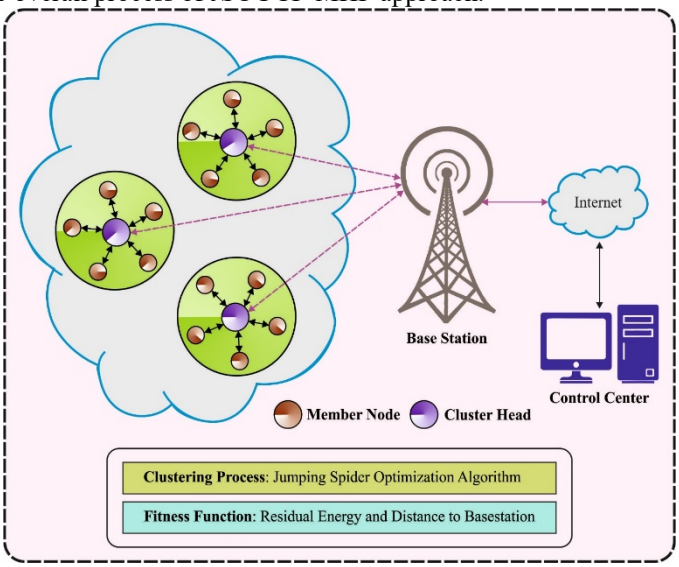



Fig. 1. Overall process of JSOUCP-MHP approach

### A. System Model

A WSN comprises $N$ uniform distribution sensors. There is one BS interconnected with users through the Internet. Let assume $S = \{S_1, S_{2,\dots}, s_i, s_{N-1}, s_N\}$ as the group of nodes, whereby $s_i$ denotes a nodei and $|S| = N$. WNS makes use of wireless radio transceiver depending on the different variables, for example, energy utilization and distance. The distance among the receiver and transmitter followed by attenuated transceiver power exponentially reduced with increased distance. Now, we represent $(x_i, \cdot\, y_i)$ and $(x_j, \cdot\, y_j)$ were the coordinates of nodes $i$ and $j$. Assume node $i$ sent to destination node $j$, the energy consumed on communication over $d$ distance can be evaluated as follows:

$$E_{\mathrm{T}x_j(l,d)} = E_{T_{X-elec}}(l) + E_{\mathrm{T}x-amp}(l,d) = \begin{cases} I \times Eelec + \varepsilon_{fs} \times d^2, d < d_0 \\ I \times Eelec + \varepsilon_{mp} \times d^4, d \geq d_0 \end{cases} \quad (1)$$

In Eq. (1), E refers to the energy consumption; Tx is to transfer, elect, amp specifies electronic and amplify to modulate, digitalized coding, spreading, and filter signals. It illustrates

that the multi-hop transmission is highly efficient in WSN. $d_o$ denotes a threshold of space model,

$$d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}}, \tag{2}$$

In Eq. (2), the power loss $\varepsilon_{fs/}\varepsilon_{mp}$ are free space and multi-path model and it is formulated by

$$E_{Rxij(l)} = E_{R_{X-elec}}(l) = l \times E_{elec}. \tag{3}$$

WSN considered $N$ node in the region of 2D $M^2$ with $k$ cluster. In relation to the hotspot perplexity in WSN was balancing the load amongst CHs according to the clustering technique. Unequal clustering decreases the cluster size nearer to BS and the size of cluster rises as distance among the BS ad well as CH rises. The cluster member senses the real-time parameter and transfers the sensed value to CH. CH aggregates and receives information to eliminate inessential information and transfer aggregated information to BS directly or through intermediatory CH. In equal clustering, the cluster size remains unchanged all over the network. But, in unequal clustering, the cluster size can be described according to the DBS [17]:

$$E_{c1uster} = E_{CH} + \left(\frac{N}{k} - 1\right) \times E_{member,} \tag{4}$$

Let $N/k - 1$ be the average of member nodes from cluster,· $E_{CH}$, $E_{members}$ denotes dissipated energy for CH and member correspondingly and it is evaluated by:

$$E_{member} = l \times E_{e1ec} + l \times \varepsilon_{fs} \times d_{toCH}^2, \tag{5}$$

$$E_{CH} = l \times E_{e1ec} \times (\tfrac{N}{k} - 1) + l \times E_{DA} \times \tfrac{N}{k} + l \times E_{e1ec} + l \times \varepsilon_{mp} \times d_{toBS}^4. \tag{6}$$

As a result, the energy consumed for WSN from the generation formulated from energy process and transceiver was given by:

$$E_{total} = E_p + E_{frame} = E_p + k \times E_{c1uster}, \tag{7}$$

In Eq. (7), $E_p$ is energy consumed for microcontroller and source voltage of node. It doesn't affect the optimization process. As a result, the consumed energy $E_{frame}$ optimizes according to the distance for clustering optimization.

### *B. Design of JSO Algorithm*

The mathematical modelling of jumping spider's hunting strategy is initially proposed. Then, The JSOA approach was introduced. The hunting strategies are searched, jumping on the prey, and attacking by persecution [18]. Also, the algorithm describes the pheromone rate of the spider as follows

Once the spider is away from a distance whereas it could catch the prey by jumping, it moves closer by indulging certain stealthy movement till it was at an attainable distance where it could catch the prey and pounces on it. The persecution approach can be described as the uniform accelerated rectilinear motion.

$$x_i = \frac{1}{2}at^2 + v_0 t \tag{8}$$

In Eq. (8), $x_i$ demonstrates the location of $i^{th}$ follower spider, $t$ refers to the time, $v_0$ represent the speed initially. The acceleration can be shown as $a = \frac{v}{t}$, where $v = \chi - \chi_0$.

In this study, to optimize, every iteration can be regarded as the duration, in which the differences between iterations are equivalent to 1 and initially the speed is fixed as zero, $v_0 = 0$. And it is formulated as follows:

$$\chi_{i(g+1)=} \frac{1}{2}\left(\overrightarrow{x_i}(g) - x_r \rightarrow (g)\right) \tag{9}$$

In Eq. (9), $\overrightarrow{x_i}(g+1)$ represents the novel location of search agent to generation $g+1$, $\overrightarrow{x_i}(g)$ indicates the present $i\text{-}th$ searching agent in generation $g$, and $\overrightarrow{x_r}(g) \rightarrow$ denotes the $r\text{-}th$ searching agent arbitrarily chosen, with $i \neq r$, whereas $r$ represents an arbitrary value ranging from 1 to the size of maximal searching agent.

The jumping spider follows the prey and jumps on them. The hunting strategy of jumping on prey is characterized by a projectile motion.

$$\begin{aligned} \overrightarrow{\chi_i} &= v_0 \cos(\alpha)\, \overrightarrow{ti} \\ \frac{dx}{dt} &= \overrightarrow{V_x} = v_0 \cos(\alpha)\overrightarrow{\imath} \end{aligned} \tag{10}$$

Likewise, vertical axis and the derivative w.r.t time can be given as.

$$\begin{aligned} \overrightarrow{y_i} &= (v_0 sen(\alpha) t_2^1 - gt^2)\vec{j} \\ \frac{dy}{dt} &= \overrightarrow{V_y} = (v_0 sen(\alpha) - gt)\vec{j} \end{aligned} \tag{11}$$

Similarly, time is characterized by strategy 1. Thus we attain the trajectory equation as follows.

$$y = x \tan(\alpha) - \frac{\overrightarrow{gx_i^2}}{2V_0^2 \cos^2(\alpha)} \tag{12}$$

Lastly, the trajectory is formulated by:

$$\overrightarrow{x_i}(g+1) = \overrightarrow{x_i}(g)\tan(\alpha) - \frac{\overrightarrow{gx^2}(g)}{2V_0^2 \cos^2(\alpha)} \tag{13}$$

$$\alpha = \frac{\varphi\pi}{180}$$

In Eq. (13), $\overrightarrow{x_i}(g+1)$ refers to the novel location of a searching agent, representing jumping spider movement, $\overrightarrow{x_i}(g)$ denotes the existing $i\text{-}th$ searching agent, $V_0$ is fixed as 100 mm/seg, $g$ denotes the gravity ($9.80665\ m/s^2$), and the $\alpha$ angle can be measured by an $\varphi$ angle arbitrarily produced within zero and one.

The Jumping spider implements a random search around the environment for locating prey. Local and global search are the two mathematical functions presented as follows.

The local search can be defined as follows:

$$\overrightarrow{x_i}(g+1) = x_{best}(g) + walk \rightarrow (\frac{1}{2} - \varepsilon) \tag{14}$$

In Eq. (14), $\overrightarrow{x_i}(g+1)$ refers to novel location of searching agent, $\overrightarrow{x_{best}}(g)$ denotes the optimal searching agent found from the preceding iteration, walk was a normally distributed pseudo-random integer within $(-2,2)$, $\varepsilon$ denotes a pseudo-random number uniformly distributing from $(0,1)$.

At the same time, the Global search can be expressed as follows.

$$\overrightarrow{x_i}(g+1) = \overrightarrow{x_{best}}(g) + (\overrightarrow{x_{best}}(g) - \overrightarrow{x_{worst}}(g))\lambda \quad (15)$$

In Eq. (15), $\overrightarrow{x_i}(g+1)$ denotes the novel location of a searching agent, $\overrightarrow{x_{best}}(g)$ and $\overrightarrow{x_{worst}}(g)$ indicates the best and worst searching agent form preceding iteration, correspondingly, and $\lambda$ refers to a Cauchy random integer with $\mu$ fixed as 0 and $\theta$ fixed as 1.

Pheromone is released by several animals, amongst which were insects, together with spiders. Nonetheless, they produce pheromone; the modelling of rate of pheromone was taken and described as follows:

$$pheromone\ (i) = \frac{Fitness_{\max} - Fitness(i)}{Fitness_{\max} - Fitness_{\min}} \quad (16)$$

In Eq. (16), $Fitness_{\max}$ and $Fitness_{min}$ represent worst and the best FV in present generation, correspondingly, while Fitness(i) refers to the present FV of $i\text{-}th$ searching agent. Eq. (16) normalizes the FV within zero and one whereas 0 indicates the worst pheromone rate, while 1 denotes the optimal. The criteria involves for lower pheromone rate value equivalent or lesser than 0.3 as follows:

$$\overrightarrow{x_i}(g) = \overrightarrow{x_{best}}(g) + \frac{1}{2}(\overrightarrow{x_{r_1}}(g) - (-1)^{\sigma_{\overrightarrow{x_{r_2}}}}(g)) \quad (17)$$

In Eq. (17), $\overrightarrow{x_i}(g)$ represent the searching agent (jumping spider) with lower pheromone rate that is upgraded, $r_1$ and $r_2$ indicates arbitrary number produced from 1 to the maximal size of searching agent, with $r_1 \neq r_2$, while $\overrightarrow{x_{r_1}}(g)$ and $\overrightarrow{x_{r_2}}(g)$ are the $r_1$, $r_2 th$ searching agent selected, $\overrightarrow{x_{best}}(g)$ indicates the best searching agent from preceding iteration and $\sigma$ indicates a binary number produced, $\sigma \in \{0,1\}$.

## III. Performance Validation

A brief set of experimental analyses is carried out to demonstrate the enhanced performance of the JSOUCP-MHP model on WSN.

Tab. 1 and Fig. 2 showcase the network lifetime (NLT) inspection of the JSOUCP-MHP model with existing models under varying density of sensor nodes (DSN). The experimental outcomes depicted that the JSOUCP-MHP model has shown enhanced performance with higher NLT values. For instance, with 50DSN, the JSOUCP-MHP model has exhibited increased NLT of 1757 whereas the enhanced metaheuristic-driven energy-aware cluster-based routing (IMD-EACBR), sunflower optimization (SFO), grey wolf optimization (GWO), and genetic algorithm (GA) models have demonstrated reduced NLT of 1549, 1433, 1422, and 1343 respectively. Moreover, with 250DSN, the JSOUCP-MHP model has attained higher NLT of 2556 whereas the IMD-EACBR, SFO, GWO, and GA models have obtained lower NLT of 2374, 2276, 2012, and 1852 respectively. On the other hand, with 500DSN, the JSOUCP-MHP model has illustrated improved NLT of 3802 whereas the IMD-EACBR, SFO, GWO, and GA models have depicted reduced NLT of 3789, 3691, 3650, and 3369 respectively.

TABLE I NLT ANALYSIS OF JSOUCP-MHP APPROACH WITH RECENT ALGORITHMS UNDER DISTINCT DSNS

| Network Lifetime (Rounds) | | | | | |
|---|---|---|---|---|---|
| Density of Sensor Nodes | JSOUCP-MHP | IMD-EACBR | SFO Algorithm | GWO Algorithm | Genetic Algorithm |
| 50 | 1757 | 1549 | 1433 | 1422 | 1343 |
| 100 | 1923 | 1789 | 1679 | 1504 | 1416 |
| 150 | 2160 | 2038 | 1773 | 1733 | 1611 |
| 200 | 2530 | 2171 | 2133 | 1870 | 1668 |
| 250 | 2556 | 2374 | 2276 | 2012 | 1852 |
| 300 | 2832 | 2666 | 2599 | 2468 | 2062 |
| 350 | 3104 | 2967 | 2852 | 2715 | 2354 |
| 400 | 3282 | 3145 | 3059 | 3050 | 2613 |
| 450 | 3487 | 3452 | 3418 | 3340 | 2976 |
| 500 | 3802 | 3789 | 3691 | 3650 | 3369 |

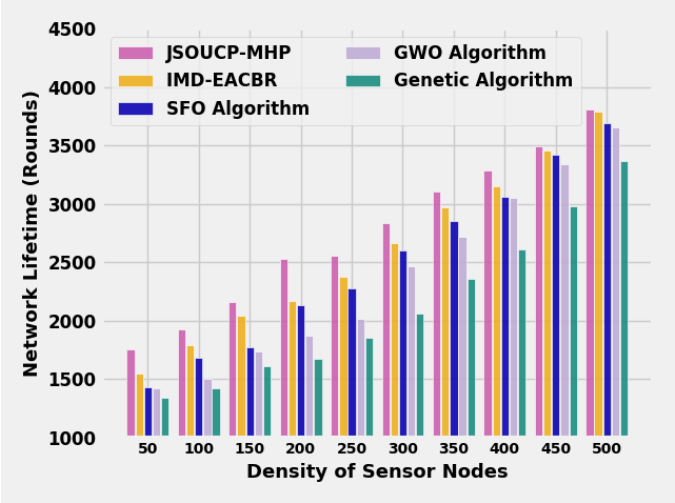



Fig. 2. NLT analysis of JSOUCP-MHP approach under distinct DSNs

## IV. Conclusions

Two In this article, a novel JSOUCP-MHP system has been presented for resolving hot spot issues in WSN. The JSO algorithm is stimulated by the characteristics of spiders naturally and mathematically modelled the hunting mechanism such as jumping skills, persecution, and search, to attack prey. The presented JSOUCP-MHP technique mainly resolves the hot spot issue for maximizing the network lifetime. To attain this, the JSOUCP-MHP technique elects proper set of CHs using average RE. Also, the JSOUCP-MHP system determines the cluster sizes dependent upon two measures such as RE and DBS. The proposed JSOUCP-MHP technique is examined under several experiments to ensure its supremacy. The comparison study shows the significance of the JSOUCP-MHP technique over other models. In upcoming years, the performance of the JSOUCP-MHP system will be improved by data aggregation approaches.

## Acknowledgment

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) program(IITP-2022-2020-0-01832) supervised by the IITP(Institute of Information &

Communications Technology Planning & Evaluation) and the Soonchunhyang University Research Fund.

## REFERENCES

[1] N. Moussa and A. E. B. E. Alaoui, "An energy-efficient cluster-based routing protocol using unequal clustering and improved ACO techniques for WSNs," Peer to Peer Networking Applications, vol. 14, no. 3, pp. 1334–1347, 2021.

[2] S. Arjunan and P. Sujatha, "A survey on unequal clustering protocols in wireless sensor networks," Journal of King Saud University - Computer and Information Sciences, vol. 33, no. 1, pp. 118, 2021.

[3] A. A. Jasim, M. Y. I. Idris, S. R. B. Azzuhri, N. R. Issa, M. T. Rahman et al. "Energy-efficient wireless sensor network with an unequal clustering protocol based on a balanced energy method (EEUCB)," Sensors, vol. 21, no. 3, pp. 784, 2021.

[4] K. W. Al-ani, F. B. Abdullah and S. Yossuf, "Unequal clustering in wireless sensor network: a review," Indonesian Journal of Electrical Engineering and Computer Science, vol. 22, no. 1, pp. 419, 2021.

[5] D. Agrawal and S. Pandey, "Load balanced fuzzy-based unequal clustering for wireless sensor networks assisted Internet of Things," Engineering Reports, vol. 2, no. 3, pp. e12130, 2020.

[6] R. Vinodhini and C. Gomathy, "Fuzzy based unequal clustering and context-aware routing based on glow-worm swarm optimization in wireless sensor networks: forest fire detection," Wireless Personal Communications, vol. 118, no. 4, pp. 3501–3522, 2021.

[7] S. Phoemphon, C. So-In, P. Aimtongkham and T. G. Nguyen, "An energy-efficient fuzzy-based scheme for unequal multihop clustering in wireless sensor networks," Journal of Ambient Intelligence and Humanized Computing, vol. 12, no. 1, pp. 873–2021.

[8] F. Zhu and J. Wei, "An energy-efficient unequal clustering routing protocol for wireless sensor networks," International Journal of Distributed Sensor Networks, vol. 15, no. 9, p. 155014771987938, 2019.

[9] N. M. Shagari, M. Y. I. Idris, R. B. Salleh, I. Ahmedy, G. Murtaza et al. "A hybridization strategy using equal and unequal clustering schemes to mitigate idle listening for lifetime maximization of wireless sensor network," Wireless Networks, vol. 27, no. 4, pp. 2641–2670,

[10] B. M. Sahoo and T. Amgoth, "An improved bat algorithm for unequal clustering in heterogeneous wireless sensor networks," SN Computer Science, vol. 2, no. 4, pp. 290, 2021.

[11] K. S. Arikumar, V. Natarajan and S. C. Satapathy, "EELTM: An energy efficient lifetime maximization approach for wsn by pso and fuzzy-based unequal clustering," Arabian Journal for Science and Engineering, vol. 45, no. 12, pp. 10245–10260, 2020.

[12] P. S. Mehra, "E-FUCA: enhancement in fuzzy unequal clustering and routing for sustainable wireless sensor network," Complex & Intelligent Systems, vol. 8, no. 1, pp. 393–412, 2022.

[13] F. Wang and H. Hu, "An energy-efficient unequal clustering routing algorithm for wireless sensor network," Revue d'IntelligenceArtificielle, vol. 33, no. 3, pp. 249–254, 2019.

[14] T. T. Nguyen, J. S. Pan and T. K. Dao, "A compact bat algorithm for unequal clustering in wireless sensor networks," Applied Sciences, vol. 9, no. 10, pp. 1973, 2019.

[15] P. C. S. Rao, P. Lalwani, H. Banka and G. S. N. Rao, "Competitive swarm optimization based unequal clustering and routing algorithms (CSO-UCRA) for wireless sensor networks," Multimedia Tools and Applications, vol. 80, no. 17, pp. 26093–26119, 2021.

[16] K. Guleria and A. K. Verma, "Meta-heuristic ant colony optimization based unequal clustering for wireless sensor network," Wireless Personal Communications, vol. 105, no. 3, pp. 891–911, 2019.

[17] T. T. Nguyen, J. S. Pan, T. K. Dao and S. C. Chu, "Load balancing for mitigating hotspot problem in wireless sensor network based on enhanced diversity pollen," Journal of Information and Telecommunication, vol. 2, no. 1, pp. 91–106, 2018.

[18] H. P. Vázquez, A. P. Delgado, P. Ranjan, C. Barde, A. Choubey et al. "A bio-inspired method for mathematical optimization inspired by arachnida salticidade," Mathematics, vol. 10, no. 1, pp. 102, 2021.

# Robust Deep Learning Model for Black Fungus Detection Based on Gabor Filter and Transfer Learning

Esraa Hassan [1]. Fatma M. Talaat[1]. Samah Adel[2]. Samir Abdelrazek[3]. Ahsan Aziz[4], Yunyoung Nam[4*], Nora El-Rashidy[1]

[1]*Faculty of Artificial Intelligence, Kafrelsheikh University, Egypt*
[2]*Departement Of Computer Eng. & Systems, Faculty of Engineering, Mansoura University, Mansoura, Egypt*
[3]*Departement of Information Systems, Faculty of Computers and Information, Mansoura University, Mansoura, 35516, Egypt.*
[4]*Department of ICT Convergence, Soonchunhyang University, South Korea*

**Contact: Yunyoung Nam (ynam@sch.ac.kr)*

*Abstract*—**Black fungus is one of the rare and dangerous mycologies that usually affects the brain and lungs and could be life-threatening in diabetic cases. Recently, some COVID-19 survivors, especially those who have co-morbid diseases, have been susceptible to black fungus. Therefore, recovered COVID-19 patients should seek medical support as soon as they notice any mucormycosis symptoms. This paper proposes a novel ensemble deep learning model that includes three different pre-trained models, including Resnet (50), VGG (19), and Inception. Our approach is medically intuitive and efficient compared to the traditional deep learning models. An image dataset was aggregated from various resources and divided into two classes, including a black fungus class and a skin infection class. To the best of our knowledge, our study is the first that is concerned with building black fungus detection models based on deep learning algorithms. The proposed approach can significantly improve the performance of the classification task and increase the generalisation ability of such a binary classification task. According to the reported results, it has been empirically achieved a sensitivity value of 0.9907, a specificity value of 0.9938, a precision value of 0.9938, and a negative predictive value of 0.9907%.**

## I. INTRODUCTION

Mucormycosis is a fatal infection that is responsible for 9–11% of fungal infections[1]. It is formalisedusually caused by direct contact with spores in the surrounding environment (i.e., soil, air, etc.). Five types of mucormycosis have been identified, including rhinocerbral (45%- 50%), cutaneous (9%- 15%), pulmonary (10%-11%), and gastrointestinal (2%-11%). Clinical infection with mucormycosis progresses rapidly, causing death in a few days. Microscopically, it shows an inflammatory change in addition to areas of necrosis. This is due to its ability to fungi into blood vessels and into various organs such as brains and could be life threatening.



It primarily affects people who have a weakened immune system, such as people who take such medications to fight germs. Therefore, Early diagnosis is considered mandatory to save patients' lives. Recently, the Indian public reported various death cases from mucormycosis (black fungus) among patients who recently recovered from COVID_19[2]. Most cases occurred among diabetic patients with uncontrolled sugar levels, and usually occurred about two weeks after COVID-19 treatment. Some medical experts attribute this to the overuse of steroids that are used to treat COVID-194678[3]–[5]. Others attribute this to the virus's nature as an immunosuppressive factor that aids in the spread of the fungus. Black fungus affects patients' skin, resulting in the appearance of black spots in different areas of the patient's face and body. Unfortunately, this effect is interleaved with other skin infections such as poikiloderma of civatte, melasma, linea nigra, etc. Therefore, the early diagnosis of black fungus using the visual traditional technique is considered a challenge. Therefore, in this paper, we tried to overcome this issue by developing a robust model for black fungus classification. It is designed to distinguish black fungus from other skin infections [6].

Image classification is considered one of the main tasks in the computer vision domain and is widely used in various domains such as object detection, age recognition, etc. In the medical domain, Image classification is utilised to categorise images to help doctors with diagnosis or another further research. The main issue in image classification is the availability of good and sufficient data, which contributes to extracting the important features in a way that permits building a trusted and robust model with good generalisation ability. Unfortunately, most medical datasets have several challenges, including size limitation, noise, and unbalanced datasets. All these factors negatively affect Deep Learning (DL) model performance. The concept of Transfer Learning(TL) was developed to resolve this issue by using pretrained models (i.e., Resnet, Inception, Xception, VGG, alexnet, etc.) that had previously been trained on huge datasets (i.e., Imagenet). Such

pretrained models help and improve classification results, especially with limited datasets.

Ensemble models are used to improve the classification accuracy and build models with lower bias and variance. The main hypothesis is that combining several weak learners correctly can result in a more robust and accurate model. Ensemble learning has two main types: homogeneous learning and heterogeneous learning. Homogeneous learning includes utilising the same classifier on different parts of the dataset. Heterogeneous learning includes using different classifiers on the same training set or different classifiers on different parts of the training set. The key points in developing an effective ensemble model are the current methods for aggregating all models. If the chosen models have low bias but high variance, the aggregated methods should tend to decrease variance. Recently, ensemble leaning has been used in several medical domains [7]. For example, Liang et al. [8]proved the effectiveness of bagging ensemble classifiers when comparing the performance with traditional learning in a medical dataset. The authors reported that bagging with the DT classifier achieves superior performance with highly imbalanced data. El-Rashidy et al. [9]propose a stacking ensemble model for predicting mortality inside the ICU. The reported performance proved the effectiveness of the proposed ensemble model [10].

Our paper's main contribution can be summarised as follows: (i) proposing an ensemble model that can be used to predict black fungus and distinguish it from other skin infections; As far as we are aware, there is no published research that could differentiate between black function and other skin illnesses. (ii) developing an ensemble model for predicting black fungus and distinguishing it from other skin infections. (iii) utilising the Gabor filter to extract the most important features. (iv) Using different pre-trained models to design our model results in an accurate result with great generalisation capability. (v) To assure and validate the suggested framework and the findings, further investigation of the impacts of utilising various optimizers, such as ADAM and SGD, under the strict supervision of medical professionals is required. (vi) In statistical comparisons with other pre-trained models, the suggested model outperformed them all. The EBF model demonstrated 0.9860% accuracy and a 0.9860% F-score, demonstrating that our model is resistant to overfitting. The remaining sections are arranged as follows: The related works are discussed in Section 2. Techniques and introductions are discussed in Section 3. Section 4 provides specifics on the suggested framework. The results and discussion were presented in Section 5, and Section 6 concluded the paper.

## II. RELATED WORK

DL is a popular area in artificial intelligence with promising results that have recently gained popularity in automatic diagnosis, classification, and segmentation systems in different domains. The existence of the COVID-19 pandemic and its consequences has rapidly increased the need for AI models. Therefore, it must be fast and accurate to provide timely support to patients as well as medical experts. Recently, several radiology images have been utilized for disease detection. For example, Tahir et al. [11] proposed a visual sensor model that could be used to mark fungal bacteria as an area of interest. They developed a AL model for fungus detection that achieved an accuracy of 94.8%. Zieliskii et al. [12] suggested a DL model dependent on deep neural networks (DNN) model for microscopic image classification of different fungi cases. Kim et al. [13]suggested a model based on Resnet(50) pre-trained model that is used for building AI models for fungus diagnosis. Their model was shown by a validation dataset using different metrics values to reach the best accuracy with an Area Under Curve (AUC) of 0.979%. This. Liu et al. [13], suggested a DNN model for applying various skin diagnoses with 16,114 cases. The DNN model separates between 26 skin cancer cases, representing 80% of COVID-19 cases, while also providing a next prediction covering 419 patients. In their model for the melanoma challenge, CNN models with varied backbones and input sizes were combined. With a few input examples, the majority of which are image-only models. Qishen Ha et al. [14] present a solid validation model, a carefully selected model target, a finely tuned pipeline, and an ensemble of various models. The cross validation AUC for their CNN model was 0.9600%. Quande Liu et al. [15] utilized two medical datasets to assess their method, including skin lesion diagnosis with the ISIC 2018 challenge and thoracic disease classification with chest x-ray. In both single-label and multi-label image classification scenarios, their technique outperforms numerous robust semi-supervised methods.

Uysal et al.[16] presented the same in to assist physicians by identifying shoulder pictures acquired from X-ray images. The performance of 26 deep learning-based pretrained models on medical dataset VGG (19), Inception (v3), and Resnet were utilized as the pre-trained models employed (50). Ensemble models were utilized with pre-trained models to achieve a robust accuracy of 0.8455, 0.8472, ROC and (AUC) of 0.8862 percent and 0.8695 percent, respectively. Rostami et al. [17] suggested an ensemble CNN to categorize surgical, diabetic, and venous ulcers in wound pictures into multi-classes. They achieved maximum and average classification accuracy for binary classification tasks of 96.4 and 94.28 percent, respectively, and 91.9 and 87.7 percent, respectively, for 3-class classification tasks. The results demonstrate that their suggested technique can be successfully built as a decision support system for clinical applications such as the classification of wound images. According to one or more pre-trained models by an individual, the prior and traditional procedures share several difficulties. In some methods, the average performance is not satisfactory. The impact on our suggested model and achieving strong accuracy for black fungus infection depends on the ensemble technique.

## III. TECHNIQUES AND PRELIMINARIES

This section introduces models and techniques that are utilised to develop our proposed model, including (1) feature extraction, (2) transfer learning, (3) pretrained models, and (4) ensemble learning.

### *A. Feature extraction technique*

The Gabor filter is a convolution filter that contains sinusoidal and Gaussian terms. It considers the optimum features that give the opportunity for multi-resolution analysis to extract the optimal features in both the frequency and spatial domains. The gaussian component is used to generate the

weights, where the sine component gives the directionality. In this paper we used a 2D Gabor filter that could be represented with the following equations:

$$g(X,y,\lambda,\theta,\psi,\sigma.\gamma) = \text{Exp}\left(\frac{X'^2+\gamma^2y'^2}{2\sigma^2}\right)\text{Exp}\left(i\left(2\pi\frac{x'}{\lambda}+\psi\right)\right) \quad (1)$$

Where $x' = xcos\,\theta + ysin\theta$ And $y' = -xsin\,\theta + ycos\theta$

Where λ is the wavelength of the component, θ is the orientation of the filter, ψ is the phase offset, σ is the standard deviation of the gaussian kernel, and γ is the spatial aspect ratio.

### B. Transfer learning approaches

In TL, the model that is used to train one task is reused for another related task after a little enhancement curated toward the specific required task. It is used for specific learning tasks where sufficient labelling images are not presented, such as in the medical domain. Transfer learning is considered a special case of deep neural networks that have many parameters that need to be trained. By using TL, the model starts with initial weights that essential just a little modification. A pre-trained model is used for new tasks in two main ways [18]. The classifier is trained on top of the pre-trained model to carry out the classification task after the pre-trained model is treated as a feature extractor [19]–[21]. The required new task was then adjusted by the entire network. As a result, the pre-trained model weights are used as the new model's starting point and updated during training. Transfer learning is frequently employed in the medical fields to address the issue of a little amount of data. We assess the effectiveness of the three most widely used models, Resnet, VGG 16, VGG 19, and Inception. The structure of these pre-trained models and each one's significance in our suggested model are briefly described in the next part.

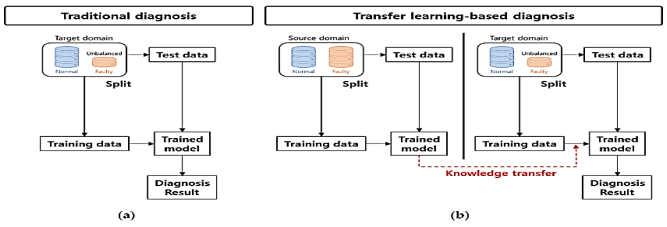



Fig. 1. The general comparison between traditional and Transfer learning approach in diagnosing [22]

### C. Inception

The inception architecture was developed by Szegedy in 2014, 30. The main goal of this module is to perform multilevel feature extractors by calculating (1 x 1), (3 x 3) and (5 x 5) convolutions, then stacking the output along the dimension before feeding to the next layer. The original version was the Google Net and its subsequent following inception. The weights of the inception model are considered small compared to VGG (19) and resnet (50). The inception model has many advantages, including (1) allowing selecting the filter size in all blocks; (2) vanishing gradients could be avoided by using several inception models and adding different objective functions for each group.

### D. VGG (19)

VGG is a network architecture that was originally introduced by Simonyan in 2014. It characterised the model's simplicity. It consists of three (3x3) convolutional layers which are stacked at increasing depths[23]. This reduction in size was handled by the max pooling layer, followed by two fully connected later, and finally the softmax classifier. VGG (16) and VGG (19) stand for the number of weights in each network. The VGG network has two main drawbacks, including (1) the network weight is large in terms of bandwidth and disk. (2) It is thought to be slow to train when compared to other networks.

### E. Resnet (50)

Resnet (50) uses exotic architecture that depends on microarchitecture, unlike CNN's traditional and sequential models (i.e., Alexnet, overfeat). The micro term refers to the building blocks that are used to build the model. It was first developed in 201532, built upon training on an ImageNet dataset.[24]

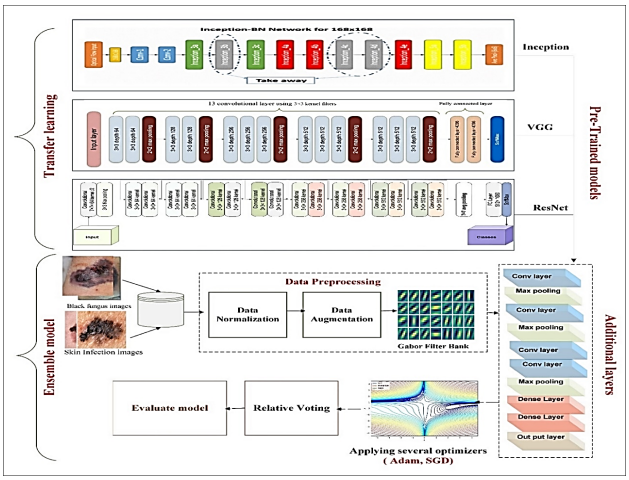



Fig. 2. General architecture of proposed ensemble model

Resnet (50) provides a smooth gradient flow which provides effective training. The core of the resnet is the "skip connection," which skips one or more layers to provide a path to the early layer and make a continuous gradient update. Resnet50 and Resnet18 are very similar in their architecture but differ in the number of layers. It is considered deeper when compared to VGG16 and VGG19. This is due to the usage of global average pooling instead of fully connected layers. It has two main advantages, including (1) models with many layers could be trained without an increase in loss percentage, and (2) it helps in tackling the vanishing gradient problem by using the identity mapping figure to show the general architecture of it.

### F. Ensemble learning

Ensemble learning is a meta-approach that used to achieve better performance by combining the predication form several

models. It proved efficiency over traditional learning. Ensemble learning widely used in deep learning, several works utilized ensemble learning to improve the overall performance of the model. This improvement returns to builds several models in a way that could reduce variance (i.e., bagging)[25], process usually used reduce bias (i.e., boosting). It also contributes to generate robust model with high generalization ability. Ensemble learning has two main types of learning. First homogeneous learning that used the same algorithms are used in different bagging, boosting) and heterogeneous learning that uses different algorithms on the same dataset (i.e., voting and stacking). The result is combined and voted from all methods in a way that prevents overfitting.[26]

## IV. The Proposed Ensemble Black Fungus (EBF)

We propose the EBF model depends on pre-trained models of the ImageNet dataset for medical imaging tasks, compared with our other proposed model, Accurate Black Fungus (ABF), which depends on three transfer learning pre-trained models. The main purpose is to propose a classification model that could overcome the challenge of the lack of training data, especially in the medical image domain [27]–[29]. The general steps used to build the classification model could be summarized as follows. (i) Loading dataset (binary class), which includes 3225 images of black fungus and uses several Gabor filters (i.e., horizontal, vertical, diagonal, etc.). To extract the important features, (ii) utilizing cross validation techniques used to choose training and testing data, (iii) integrating three different pre-trained models, including inception, VGG (19), and resent (50), are used to generate a binary classification model, (iv) for each pre-trained model. The last layer was removed and replaced with a group of CONV, Maxpooling, and dropout layers. Details of the added layers mentioned in subsection 4.3: (v) implementing an ensemble model that integrates the three pre-trained models, then using relative voting to obtain the result, and (vi) using various metrics to measure the reliability and accuracy of the proposed model. The superiority of the proposed model result ensures the ability to be utilized for any medical imaging task with a small number of labelled images. Fig. 2 shows the general architecture of the proposed model.

### A. Augmentation Techniques

In this paper, we utilized some data augmentation techniques that we employed random rotation ranges between 45 and 180 degrees, width shift range with a value of 0.2, height shift range with a value of 0.2, applying horizontal flip and vertical flip. As we mentioned above, we used the Gabor filter in extracting the optimal features. Numerous digital features (kernels) can be identified by varying the Gabor filter parameters [30]. For example, changing the theta value to zero will generate horizontal features, while making it generate vertical features will not. One filter is usually not enough to extract the important features [31][32]. Therefore, tens of filters are created and then passed to the AFB model. The following Table 1 includes some examples of the generated features.

TABLE I. THE VALUES AND SIZES OF GABOR FILTER PARAMETERS

| Num/Paramter | Ksi | Lamda($\lambda$) | $Theta($ | Phi($\psi$) | $Sigma($ ) | $gamma($ | ze |
|---|---|---|---|---|---|---|---|
| 1 | 15 | 1*np.pi/4 | 1*np.pi/4 | 0.8 | 5 | 0.1 | |
| 2 | 15 | 1*np.pi/4 | 1*np.pi/2 | 0.8 | 5 | 0.1 | |
| 3 | 15 | 1*np.pi/4 | 1*np.pi/4 | 1 | 5 | 0 | |
| 4 | 9 | 1*np.pi/4 | 1*np.pi/2 | 1 | 10 | 0.99 | |
| 5 | 15 | 1*np.pi/4 | 1*np.pi/4 | 0 | 0 | 0.1 | |

### B. The classification model (AFB)

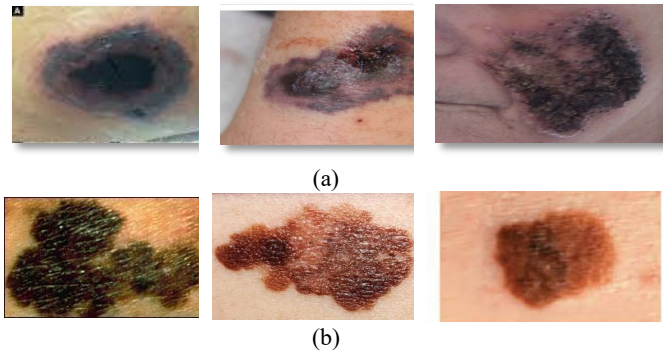
The AFB model is based on an efficient DCNN model that integrates many advanced elements to address a change of issues, including better feature extraction. In Table 2, a thorough explanation of the suggested model is provided. The input of the suggested architecture in the instance of the black fungus classification models is (165 154). the changed input size is (80 80). Two conventional convolutional layers that operate in succession are the model's initial stages. The first one has a kernel size of 2x2 and a filter number of 128. The kernel size of the second convolution is 2x2. ReLu layers come after both convolutional layers. Finally, one dropout layer is used to join fully connected layers. The result is finalized using SoftMax.

## V. Results and Discussion

The experimental setup, the preparation procedures, the dataset used, the assessment metrics, and the outcomes of our suggested strategy are all included in this part. The performance of our suggested model was compared to the performance of each pre-trained model separately to ensure its efficacy.

### A. Dataset

The Black Fungus and Melanoma (BFM) dataset [33] contains three compressed folders with images inside. The first file contains 260 images of black fungus after applying data preprocessing, which becomes 1300 images. The second file contains 385 images of melanoma after applying data preprocessing, which becomes 1925. The BFM dataset contains images, as shown in Figure 3, that facilitate training and validation during the utilisation of deep learning algorithms for skin disease recognition and classification.



(a)

(b)

Fig. 3. (a) Samples of black fungus cases for skin infection, (b) Samples of Melanoma cases for skin infection

## VI. Conclusions

Pathologists' knowledge and experience are used to diagnose black fungus, and reports from different laboratories doing physical inspections may differ. The diagnosis of black fungus is usually dependent on a pathologist's knowledge and expertise as well as laboratory reports, which may vary from one to another. The EBF model provides automatic classification for black fungus images. This paper presents the EBF model for classifying black fungus images, which has proven to be a useful and promising method for detection and categorizing information from cases. The results are compared to both traditional machine learning methods and CNN-based approaches, proving that deep learning-based image categorization is feasible. To accomplish this, the dataset conversion rate must be increased. Without the requirement for manual feature extraction, it can carry out binary and multi-class classification. As a result, it is regarded as a promising and useful technique for diagnosing black fungus. The outcomes combine three widely used pre-trained models: VGG (19), Resnet (50), and Inception (v3). It achieved values of 0.9907% for sensitivity, 0.9938 for specificity, 0.9938 for precision, 0.9907% for negative predictive value, 0.0062% for false positive rate, 0.0062% for false discovery rate, 0.0093% for false discovery rate, 0.9922% for accuracy, 0.9922% for F-score, and 0.9845% for Matthews Correlation Coefficient.

## Acknowledgment

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) program(IITP-2022-2020-0-01832) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation) and the Soonchunhyang University Research Fund.

## References

[1] Y. Liu et al., "A deep learning system for differential diagnosis of skin diseases," Nat. Med., vol. 26, no. 6, pp. 900–908, 2020, doi: 10.1038/s41591-020-0842-3.

[2] N. Hameed, A. M. Shabut, M. K. Ghosh, and M. A. Hossain, "Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques," Expert Syst. Appl., vol. 141, p. 112961, 2020, doi: 10.1016/j.eswa.2019.112961.

[3] J. Cafasso, "Is It Mild Cognitive Impairment or Something Else?," Healthline, no. Mci, pp. 1–11, 2020, [Online]. Available: https://www.healthline.com/health/is-it-mild-cognitive-impairment-or-something-else#MCI-vs.-dementia-vs.-healthy-aging

[4] T. Akter et al., "Improved Transfer-Learning-Based Facial Recognition Framework to Detect Autistic Children at an Early Stage.," Brain Sci., vol. 11, no. 6, May 2021, doi: 10.3390/brainsci11060734.

[5] X. Bai et al., "Development and Evaluation of a Machine Learning Prediction Model for Small-for-Gestational-Age Births in Women Exposed to Radiation before Pregnancy," J. Pers. Med., vol. 12, no. 4, 2022, doi: 10.3390/jpm12040550.

[6] S. Albahli and G. N. A. H. Yar, "Efficient Grad-Cam-Based Model for COVID-19 Classification and Detection," Comput. Syst. Sci. Eng., vol. 44, no. 3, pp. 2743–2757, 2023, doi: 10.32604/csse.2023.024463.

[7] V. Nirmala and B. Gomathy, "Hybrid Deep Learning Method for Diagnosis of Cucurbita Leaf Diseases," Comput. Syst. Sci. Eng., vol. 44, no. 3, pp. 2585–2601, 2023, doi: 10.32604/csse.2023.027512.

[8] A. Sroka-oleksiak, D. Rymarczyk, A. Piekarczyk, and M. Brzychczy-wloch, "Deep learning approach to describe and classify fungi microscopic images," 2020.

[9] N. Mahmoud, S. El-sappagh, T. Abuhmed, and S. M. Abdelrazek, "Intensive care unit mortality prediction : An improved patient-specific stacking ensemble model," vol. XX, 2020, doi: 10.1109/ACCESS.2020.3010556.

[10] V. Gautam et al., "Early Skin Disease Identification Using Deep Neural Network," Comput. Syst. Sci. Eng., vol. 44, no. 3, pp. 2259–2275, 2023, doi: 10.32604/csse.2023.026358.

[11] [11]M. W. Tahir, N. A. Zaidi, A. A. Rao, R. Blank, M. J. Vellekoop, and W. Lang, "A Fungus Spores Dataset and a Convolutional Neural Network Based Approach for Fungus Detection.," IEEE Trans. Nanobioscience, vol. 17, no. 3, pp. 281–290, Jul. 2018, doi: 10.1109/TNB.2018.2839585.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conf. Comput. Vis. Pattern Recognit., pp. 770–778, 2016.

[13] S. H. Cho, S. Kim, and J.-H. Choi, "Transfer Learning-Based Fault Diagnosis under Data Deficiency," Appl. Sci., vol. 10, no. 21, 2020, doi: 10.3390/app10217768.

[14] L. Y. K. Nakada and R. C. Urban, "COVID-19 pandemic: environmental and social factors influencing the spread of SARS-CoV-2 in São Paulo, Brazil," Environ. Sci. Pollut. Res., vol. 28, no. 30, pp. 40322–40328, 2021, doi: 10.1007/s11356-020-10930-w.

[15] H. tao Zhang et al., "Automated detection and quantification of COVID-19 pneumonia: CT imaging analysis by a deep learning-based software," Eur. J. Nucl. Med. Mol. Imaging, vol. 47, no. 11, pp. 2525–2532, 2020, doi: 10.1007/s00259-020-04953-1.

[16] S. Armstrong, "Covid-19: Tests on students are highly inaccurate, early findings show.," BMJ, vol. 371, p. m4941, Dec. 2020, doi: 10.1136/bmj.m4941.

[17] P. Boldog, T. Tekeli, Z. Vizi, A. Dénes, F. A. Bartha, and G. Röst, "Risk Assessment of Novel Coronavirus COVID-19 Outbreaks Outside China," 2020, doi: 10.3390/jcm9020571.

[18] S. D. Mahalakshmi, "An Optimized Transfer Learning Model Based Kidney Stone Classification," Comput. Syst. Sci. Eng., vol. 44, no. 2, pp. 1387–1395, 2023, doi: 10.32604/csse.2023.027610.

[19] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," Sci. Rep., vol. 10, no. 1, pp. 1–12, 2020, doi: 10.1038/s41598-020-76550-z.

[20] S. Venkatesh and S. Gopal, "Expert Systems with Applications Robust Heteroscedastic Probabilistic Neural Network for multiple source partial discharge pattern recognition – Significance of outliers on classification capability," Expert Syst. Appl., vol. 38, no. 9, pp. 11501–11514, 2011, doi: 10.1016/j.eswa.2011.03.026.

[21] P. M. Graffy, V. Sandfort, R. M. Summers, and P. J. Pickhardt, "Automated Liver Fat Quantification at Nonenhanced Abdominal CT for Population-based Steatosis Assessment,"

Radiology, vol. 293, no. 2, pp. 334–342, 2019, doi: 10.1148/radiol.2019190512.

[22] S. H. Cho and S. Kim, “applied sciences Transfer Learning-Based Fault Diagnosis under Data Deficiency,” 2020.

[23] G. E. Atteia, “Latent Space Representational Learning of Deep Features for Acute Lymphoblastic Leukemia Diagnosis,” Comput. Syst. Sci. Eng., vol. 45, no. 1, pp. 361–376, 2023, doi: 10.32604/csse.2023.029597.

[24] S. Tao, “Deep Neural Network Ensembles,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 11943 LNCS, pp. 1–12, 2019, doi: 10.1007/978-3-030-37599-7_1.

[25] H. Liu and M. Cocea, “Nature-inspired framework of ensemble learning for collaborative classification in granular computing context,” Granul. Comput., vol. 4, no. 4, pp. 715–724, 2019, doi: 10.1007/s41066-018-0122-5.

[26] H. Saoud, A. Ghadi, and M. Ghailani, “Breast cancer diagnosis using machine learning and ensemble methods on large seer database,” J. Theor. Appl. Inf. Technol., vol. 99, no. 3, pp. 594–604, 2021.

[27] B. Sluban and N. Lavra, “Neurocomputing Relating ensemble diversity and performance : A study in class noise detection,” vol. 160, pp. 120–131, 2015, doi: 10.1016/j.neucom.2014.10.086.

[28] T. E. Mathew, K. S. A. Kumar, and K. S. Kumar, “Breast Cancer Diagnosis using Stacking and Voting Ensemble models with Bayesian Methods as Base Classifiers,” vol. IX, no. Ii, pp. 108–121, 2020.

[29] B. Peter, Bagging , Boosting and Ensemble Methods Bagging , Boosting and Ensemble Methods, no. January 2012. 2015. doi: 10.1007/978-3-642-21551-3.

[30] A. T. Lopes, E. de Aguiar, A. F. de Souza, and T. Oliveira-Santos, “Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order,” Pattern Recognit., vol. 61, pp. 610–628, 2017.

[31] D. Al-karawi, S. Al-Zaidi, N. Polus, and S. Jassim, “Machine Learning Analysis of Chest CT Scan Images as a Complementary Digital Test of Coronavirus (COVID-19) Patients,” 2020, doi: 10.1101/2020.04.13.20063479.

[32] Y. Chen, L. Zhu, P. Ghamisi, X. Jia, G. Li, and L. Tang, “Hyperspectral Images Classification With Gabor Filtering and Convolutional Neural Network,” no. March 2018, 2017, doi: 10.1109/LGRS.2017.2764915.

[33] E. Hassan, “Black fungs dataset.” https://data.mendeley.com/datasets/79hbdrfjg8

# A Novel Deep Learning-Based Model for Classification of WheatGene Expression

Amr Ismail [1], Walid Hamdy[1,2], Aya M. Al-Zoghby[3], Wael A. Awad[3], Ahmed Ismail Ebada[3], Yunyoung Nam[4], Byeong-Gwon Kang[4,*] and Mohamed Abouhawwash[5,6]

[1]*Faculty of Science, Port Said University, Port Said, Egypt*
[2]*Modern Academy for Computer Science and Management Technology, Cairo 11742, Egypt*
[3]*Faculty of Computers and Artificial intelligence, Damietta University, New Damietta, Egypt.*
[4]*Department of ICT Convergence, Soonchunhyang University, South Korea*
[5]*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.*
[6]*Department of Computational Mathematics, Science, and Engineering (CMSE), Michigan State University, East Lansing, MI, 48824 USA.*

**Contact: Byeong-Gwon Kang. Email: bgkang@sch.ac.kr*

*Abstract*—**Deep learning (DL) plays a critical role in processing and converting data into knowledge and decisions. DL technologies have been applied in a variety of applications, including image, video, and genome sequence analysis. In deep learning the most widely utilized architecture is Convolutional Neural Networks (CNN) are taught discriminatory traits in a supervised environment. In comparison to other classic neural networks, CNN makes use of a limited number of artificial neurons, therefore it is ideal for the recognition and processing of wheat gene sequences. Wheat is an essential crop of cereals for people around the world. Wheat Genotypes identification has an impact on the possible development of many countries in the agricultural sector. In quantitative genetics prediction of genetic values is a central issue. Wheat is an allohexaploid (AABBDD) with three distinct genomes. The sizes of the wheat genome are quite l large compared to many other kinds and the availability of a diversity of genetic knowledge and normal structure at breeding lines of wheat, Therefore, genome sequence approaches based on techniques of Artificial Intelligence (AI) are necessary. This paper focuses on using the Wheat genome sequence will assist wheat producers in making better use of their genetic resources and managing genetic variation in their breeding program, as well as propose a novel model based on deep learning for offering a fundamental overview of genomic prediction theory and current constraints. In this paper, the hyperparameters of the network are optimized in the CNN to decrease the requirement for manual search and enhance network performance using a new proposed model built on an optimization algorithm and Convolutional Neural Networks (CNN).**

## I. INTRODUCTION

Cultivated crops must be increased to meet the world's population's food, feed, and fuel demand projected at more than 9 billion by 2050 [1]. One in nine people currently finds themselves living under food insecurity [2]. With limited opportunities to expand farming on existing land, increasing yields could dramatically reduce the number of people at risk of starvation [3]. Given the need to increase crop production by 50 percent by 2050 [4], our current yield levels are inadequate to achieve this target [5]. Therefore, it is necessary and urgent to find ways to boost crop productivity, such as by genetically modifying cultivars and improving agricultural practices [6, 7]. The plant sector is the center of many countries' production. Growing plant typically has special features, such as habits, morphology, and economic value. According to statistics, several plants are registered and named worldwide [8]. We apply genomic prediction techniques in the plant recognition and identification study to make this industry successful. New approaches and techniques in the detection of plant diseases are being employed in the Genomic processing industry. Therefore, in recent years, researchers have become involved in the detection of plant diseases by using genomic processing technology for their importance and effect on farming's future. However, the prediction of the wheat gene is a new problem in machine learning. Through this method, the goal is to achieve a perfect model for wheat gene expression.

## II. RELATED WORK

Deep learning is developing into a strong form of machine learning, which benefits both the outstanding computational resources and the very large available datasets [9]. The need to specifically define which features to use or use for data analysis is bypassed by deep learning. Deep learning then optimizes a robust end-to-end cycle by mapping data samples to outputs compatible with the large identified network training data sets. The CNNs practice this end-to-end mapping for image processing activities, by optimizing several layers of filters. The first filters are interpreted simply as low-level image features (e.g., borders, bright spots, color variations), and the subsequent layer combinations are more and more complex. CNN greatly outperforms all current alternative methods for image analysis where adequate training is given. Results improved from 84.6 percent in 2012 [10] to 96.4 percent in 2015 [11] with

benchmark-classification tasks attempting to determine which one thousand different objects are pictures.

Machine Learning technology have be used in a lot of applications in recent years, including image processing. CNN as indicated in [12] is the most common architecture and is primarily used in deep analysis. The CNN is equipped with discriminatory learning features by supervised means. In contrast to other conventional neural networks, CNN utilizes a few artificial neurons that make it suitable for image detection and processing. On the other hand, for training phases, CNN needs a broad sample number. CNN also has hyperparameters and a wide range of special architectures that are considered expensive and difficult to identify manually such as optimum hyperparameters [13]. We are responsive to the planning, which dramatically impacts CNN efficiency, of certain hyperparameters. Moreover, the hyperparameters for each dataset have to be modified because the over-parameters are different from one dataset to another. The correct values for hyperparameters for a certain dataset are calculated by trial and error since a math format is not given to manually change the hyperparameters. Selecting hyperparameter values requires detailed data that forces non-experts to use a random search or a grid seeking to find the better hyperparameters, which achieve the best performance of CNN. In [14] They used six deep neural networks and machine learning techniques to investigate and exploit the methylation patterns of the Chinese spring bread wheat cultivar in order to identify differentially expressed genes (DEGs) between leaves and roots. Genes with increased terms at leaves were mostly engaged in pigment and photosynthesis production activities, as expected, whereas genes with no difference in expression amidst leaves and roots were mostly implicated in protein processing and diaphragm structures. In [15] They used this study to see how well the DL model worked in the spring wheat breeding programme at Washington State University. They compared and evaluated the execution of two DL techniques, the convolutional neural network (CNN) and the multilayer perceptron (MLP), ridge retraction better linear equitable predictor (rrBLUP), which is a popular GS model. They used the nested association mapping (NAM) for the Spring wheat many seeded from the 2014–2016 growth seasons yielded 650 recombinant inbred lines (RILs). They used cross-validations (CVs), alternative sets, and independent validations of SNP markers, they made predictions for five various quantitative variables using various genetic architectures. Hyperparameters for models of DL were adjusted by decreasing the root average square in the training dataset and employing dropout and regularization to avoid model overfitting.

In [16] they used R-CNN Faster to verify the spike number by using the dataset for high-density wheat 660K SNP array. they achieved an accuracy of 86.7%. They approve that the R-CNN Faster model is faster and has a high accuracy that may be applied to genetic investigations of SN in wheat.

## III. Deep Learning Principles

A standardized DL architecture consists of a mixture of multiple "neurons" layers. In the 50s, with a prominent "perceptron" of Rosenblatt, the idea of a nerve network was proposed, inspired by the activity of the brain [17]. In the past decade, the DL resurgence was focused on the development of powerful algorithms which can be used in complex network parameters containing multiple layers of neurons (e.g. backpropagation) [18] and on the fact that they surpass current algorithms in various automated recognizing functions like picture checking [19]. Deep learning is an area of many specific jargon terms, which means that some of the most crucial terms are defined in Figure 1 to make understanding easier for an inexperienced user.

Fig. 1 Multistage perceptron (MLP) graph displaying the feedback of the simple "Neuron" with n inputs and four hidden layers of single nucleotide polymorphisms (SNPs). The linear combinations' nonlinear transformations (xi, wi, and biases b) all culminate in a single neuron. where $x_i$ represents the neuron's i input, $w_i$ represents a weight connected by input i, and b represents a time-invariant alignment level.

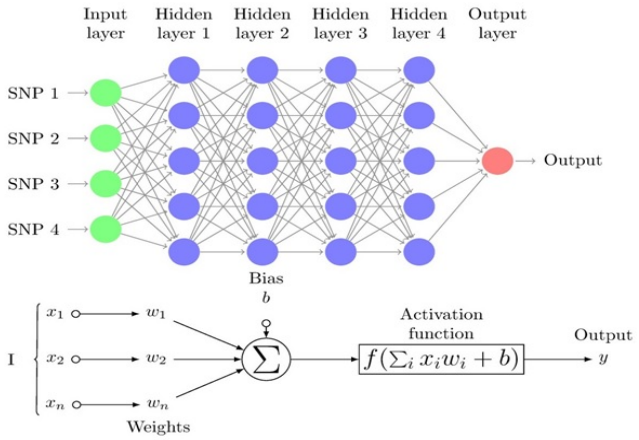



Fig. 1. Multistage perceptron (MLP).

### A. Deep Learning Architectures

Although all DL techniques generally use stacked neuron layers, they do also include a large architecture. The most prevalent ones are convolutional neuro-networks (CNN), multilayer perceptron (MLP), generative opposing networks (GANs), and recurrent neural networks (RNNs). These are listed in effect, although the reader should be aware of various additional options [20].

The multi-layer perceptron network (MLP) consists of a set of completely connected layers named hidden and input layers (see Figure 2) and is one of the most common DL architectures. The first layer receives SNP genotypes (x) feedback in the sense of genomic prediction [21], while the initial layer's output is a weighted, non-linear function of all feedback plus a "bias". Then the first output layer is shown in Eq. (1):

$$z^{(1)} = b_0 + W^{(0)} f^{(0)}(x) \quad (1)$$

When x includes each individual's genotypes, b is considered a "bias" and is measured along with the remaining weights $W^0$ and f is a nonlinear function

(activation function available on Keras). The same term is used in successive layers so that the neuron's inputs of a certain layer are the outputs of the preceding layer $z^{(k-1)}$:

$$z^{(k)} = b_k + W^{(k-1)} f^{(k-1)}(z^{(k-1)}) \quad (2)$$

The final layer generates a number matrix, whether the goal is a true phenotype, or if the goal is a class (ie a problem with classifying) an array of probabilities for each point. Although MLPs constitute a powerful strategy for managing classification or regression issues, they are not the perfect way to handle space or time sets [22]. In latest years, other methods of DL have been suggested in order to deal with these challenges, such as recurrent neural networks, deep generative networks, or convolutional neural networks.

Input variables have been spread in accordance with space models with one dimension (for example, SNPs or text) and two or three dimensions (for example, images), to conform to the circumstances of the implementation of neural networks. CNN's have been introduced. CNN is a particular type of neural network that uses convolution in hidden layers rather than of full matrix reproduction [23]. A CNN consists of thick layers and "convolutional layers" that are fully connected (Fig. 2). An overall operation as well as the input of predetermined width and steps are done in every convolutionary layer. A 'kernel' or 'buffer' is a collection of convolutional processes that functions similarly to a 'neuron' in an MLP [24].

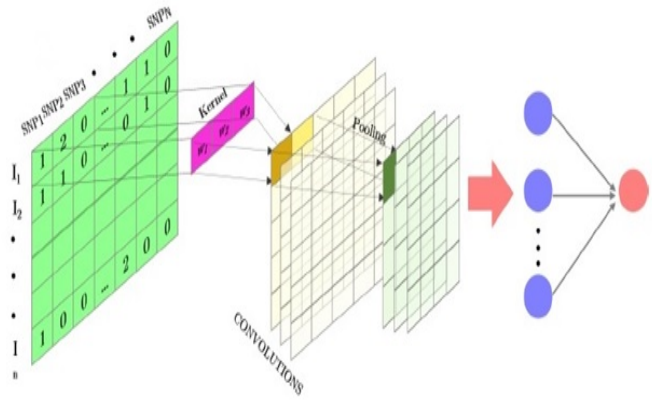



Fig. 2. Total view of 1D fully convolutional SNP-matrix neural network.

After each convolution, the output is generated using an activation function. Finally, the results are frequently evened out through a "pooling" method. The kernel outputs of the various positioning positions are combined by using all values of those positions on average, maximum, or minimum. Its capability to which the amount of parameters to be determined is one of the main advantages of convolution networks. These networks have already restricted connections and are translations similar. Fig. 3 provides an example of a one-dimensional (1D) kernel convolution with a scale of 3K [25].

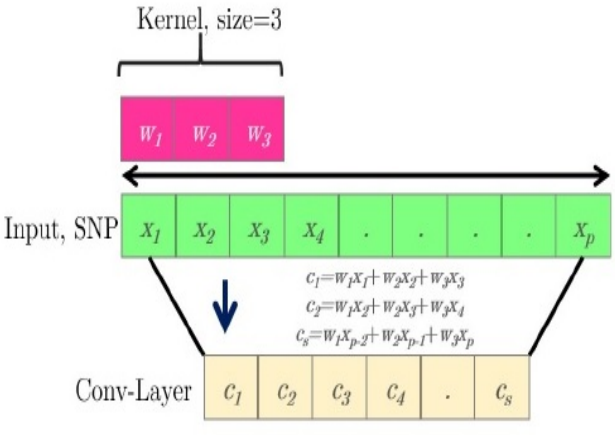



Fig. 3. Simple one-dimensional (1D) operation scheme.

### B. Convolutional Neural Network

The CNN is so good at categorizing simple patterns in data, it might be utilized to build additional complicated patterns during higher layers. CNNs are a specific type of multilayer neural network. It is trained using the backpropagation algorithm, which is used by practically all other neural networks. CNN's architecture sets it apart from the competition. In a CNN design, there are input layers, numerous hidden levels, and output layers. The hidden layer is made up of pooling layers, Convolutional layers, and fully connected layers [26].

The input data is received by the convolution layer, which applies a filter to it, essentially, the input data is multiplied by the kernel to generate the adjusted output data. A convolution layer subsampling method is the Pooling layer. The goal is to reduce the number of dimensions. An input layer serves as the first layer in the proposed CNN algorithm used in this study. The second layer makes up a one-dimensional convolution layer with three kernel sizes, a 30 filter, and RELU activation. The third layer is the max pooling layer, with two pool sizes. The next layer is a completely connected layer with the ability to activate RELU.

Finally, the output layer is made up of a single sigmoid activation in a neuron. The ADAM optimizer is applied for learning, as a cost function with binary cross-entropy.

### C. Recurrent Neural Network

RNN is the only algorithm with internal memory. Therefore, it is a very powerful and reliable algorithm, the RNN is incredibly powerful since it is still the only algorithm with internal memory. The internal memory of the RNN allows the algorithm to recall and research critical information about the input it receives; this ability allows the program to predict what will happen next with great accuracy [27]. The information in an RNN loops back on itself. As demonstrated in Fig. 4, it considers the current input as well as what it has learned from previous inputs before making a decision.

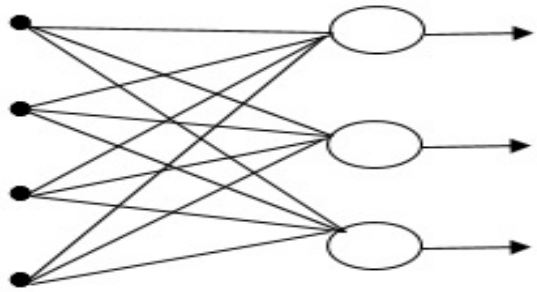


Fig. 4. Recurrent Neural Network.

This study employs a simple RNN layer, with the output being fed back into the input. The simple RNN layer is used to apply the RELU activation function. A sigmoid activation algorithm was also employed for the output layer. For learning, the ADAM optimizer is employed, and as a cost function, binary cross-entropy is used.

## IV. The Proposed Approach

In this section, the dataset used to implement the proposed approach is first described, then the details of the approach proposed are explained.

### *A. Dataset description*

The data for this study is from the Global Wheat data set, which is open to the public [28]. The original genotypic data consisted of 73,345 polymorphic markers anchored to the Chinese Spring Ref Seqv1 map. Before filtering the genotypic data, RILs with lacking phenotypic data in a single setting were deleted. SNP markers having a missing data rate of higher than 20%, minor allele frequencies of less than 0.10, and RILs with more than 10% genotypic data were also eliminated, leaving 40,000 SNP and 635 RILs markers for analysis. Using 635 RILs and 40,000 SNP markers. The demographic structure of the 26 NAM families was investigated using principal component analysis (PCA).

### *B. The Proposed Approach for Classification of Wheat Gene Expression*

The solution suggested is an incredibly effective way of optimizing the efficiency of the CNN network by the incorporation of the terminals of two pre-trained CNN networks. In fact, the model's hyperparameters are designed such that each model performs better.

## V. EXPERIMENTS AND RESULT

The protein interaction network was mapped to the gene expression levels from our dataset. Each time, 95 samples were used as testing data and 285 samples were used as training data to train the convolutional neural network architecture. There were 60 epochs in total. Then, by 95 samples (4x95=380) and k=4, we rank-fold cross-validation (CV). The selecting test data were then randomized to a 10-time process of randomization, after which the average value for the following machine learning metrics—accuracy, specificity, recall (sensitivity), and precision—was calculated. These matrices show the relationship between correctly and incorrectly predicted outcomes. TN (True Negative), TP (True Positive), FN (False Negative), and FP (False Positive) are the four categories in the confusion matrix. which were defined as follows:

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN) \quad (3)$$

$$\text{recall} = TP/(TPFN) \quad (4)$$

$$\text{precision} = TP/(TPFP) \quad (5)$$

$$\text{specificity} = TN/(TNFP) \quad (6)$$

The size and number of convolutional filters, as well as the number and size of convolutional layers and hidden layers, were all examined in various combinations. With the architecture, the best outcomes were obtained. Tab. 1 shows that with our sample, with a mean accuracy of 99.4%, the improved DNN was the most accurate, followed by DNN with 98.2% and 97.5% for CNN and RNN, respectively.

**TABLE I. COMPARISON OF CLASSIFICATION ACCURACY RESULTS WITH THE IMPROVED DNN, DNN, RNN, AND CNN.**

| epoch number | Improved DNN | DNN | RNN | CNN |
|---|---|---|---|---|
| Epoch 1 | 99.1 | 98.1 | 97.5 | 97.4 |
| Epoch 2 | 99.3 | 98.3 | 97.4 | 97.3 |
| Epoch 3 | 99.4 | 98.0 | 97.2 | 97.5 |
| Epoch 4 | 99.0 | 98.2 | 97.3 | 97.4 |
| Epoch 5 | 98.9 | 98.3 | 97.6 | 97.1 |
| Epoch 6 | 99.0 | 97.98 | 97.4 | 97.5 |
| Epoch 7 | 98.8 | 97.99 | 97.3 | 97.4 |
| Epoch 8 | 99.4 | 98.0 | 97.5 | 97.3 |
| Epoch 9 | 99.3 | 98.1 | 97.1 | 97.2 |
| Epoch 10 | 99.2 | 98.2 | 97.4 | 97.4 |

Overall, the improved DNN algorithm can be observed that attained maximum accuracy in this study's dataset. There were 100 epochs in total. Fig. 7 shows the curves of the high-accuracy model discovered by Improved DNN on the convex dataset for 100 epochs when compared to DNN, RNN, and CNN models. We can see that accuracy of our models have improvement when compared to other models, implying that the Improved DNN is actually capable of identifying perfect models for a given dataset.

## VI. Conclusion

In this paper, we have presented a novel deep learning-based model which improves DNN by applying the dropout model to classify Wheat gene expressions. In addition to, the deep learning algorithms CNN, DNN, and RNN, and the proposed model are implemented for the classification of gene expression data. Moreover, the outliers and noisy data are addressed, by using a pre-processing methodology for all features of gene expression, after that we trained all of our models individually using a perfect framework and learning method. Finally, our learned models are applied to testing data to classify it. For all of the datasets studied, the Improving-DNN outperformed other models in accuracy terms from the result illustrated our Improving-DNN has a high accuracy of 99.4%, while DNN has 98.2% accuracy, RNN and CNN have 97.5% accuracy.

Therefore, the Improving-DNN model is actually more appropriate for solving the wheat gene expression dataset.

## ACKNOWLEDGMENT

This research was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0012724, The Competency Development Program for Industry Specialist) and the Soonchunhyang University Research Fund.

## REFERENCES

[1] U.N. Desa,"World population prospects 2019: Highlights," New York (US): United Nations Department for Economic and Social Affairs, vol 11, no. 1, pp. 125. 2019.

[2] S. McGuire, "FAO, IFAD, and WFP. The state of food insecurity in the world 2015: meeting the 2015 international hunger targets: taking stock of uneven progress. Rome: FAO," Advances in Nutrition, vol. 6, no. 5, pp. 623-624, 2015.

[3] M.W.Rosegrant, S.Tokgoz, P.Bhandary and S. Msangi, "Looking Ahead: Scenarios for the Future of Food. 2012 Global Food Policy Report. IFPRI. Washington," International Food Policy Research Institute (IFPRI), vol. 4, no. 3, pp. 1-15, 2013.

[4] D. Tilman, C. Balzer, J. Hill and B. L. Befort, "Global food demand and the sustainable intensification of agriculture," Proceedings of the National Academy of Sciences - PNAS, vol. 108, no. 50, pp. 20260–20264, 2011.

[5] D. K. Ray, N. D. Mueller, P. C. West and J. A. Foley, "Yield trends are insufficient to double global crop production by 2050," PlOS ONE, vol. 8, no. 6, pp. e66428–e66428, 2013.

[6] H. Spiertz, "Avenues to meet food security. The role of agronomy on solving complexity in food production and resource use," European journal of agronomy, vol. 43, no. 5, pp. 1–8, 2012.

[7] J.L. Araus, R. Park, D. Calderini, D. Miralles, T. Shen et al., "Prospects of doubling global wheat yields," Food and energy security, vol. 2, no. 1, pp. 34–48, 2013.

[8] W.Hamdy, A. Ismail, W. A. Awad, A.H. Ibrahim, and A. Hassanien, "A Support Vector Machine Model for Rice (Oryza sativa L.) Leaf Diseases Based on Particle Swarm Optimization." In Artificial Intelligence: A Real Opportunity in the Food Industry, Springer, Cham, pp. 45-54, 2023.

[9] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," Nature (London), vol. 521, no. 7553, pp. 436–444, 2015.

[10] A. Elaraby, W. Hamdy and M. Alruwaili, "Optimization of deep learning model for plant disease detection using particle swarm optimizer," Computers, materials & continua, vol. 71, no. 2, pp. 4019-4031, 2022.

[11] J. R. Ubbens and I. Stavness, "Corrigendum: deep plant phenomics: A deep learning platform for complex plant phenotyping tasks," Frontiers in plant science, vol. 8, no. 12, pp. 2245–2245, 2018.

[12] A. Elaraby, W. Hamdy and S. Alanazi, "Classification of citrus diseases using optimization deep learning approach," Computational intelligence and neuroscience, vol. 2022, no. 10, pp. 9153207–9153212, 2022.

[13] N. Ni and S. Xu, "Model optimization strategies based on deep neural networks Learning and application of pruning optimization algorithms," Journal of Physics, Conference series, vol. 2303, no. 1, pp. 012033, 2022.

[14] A. N. Diaye, B. Byrns, A.T. Cory, K.T. Nilsen, S. Walkowiak et al., "Machine learning analyses of methylation profiles uncovers tissue-specific gene expression patterns in wheat," The plant genome, vol. 13, no. 2, pp. e20027, 2020.

[15] K. S. Sandhu, D. N. Lozada, Z. Zhang, M. O. Pumphrey and A. H. Carter, "Deep learning for predicting complex traits in spring wheat breeding program," Frontiers in plant science, vol. 11, no. 4, pp. 613325–613335, 2021.

[16] L. Li, M.A. Hassan, S. Yang, F. Jing, M. Yang et al., "Development of image-based wheat spike counter through a Faster R-CNN algorithm and application for genetic studies." The Crop Journal, vol. 12, no. 3, pp. 1-12, 2022.

[17] P. Matteo, "Machines that morph logic: neural networks and the distorted automation of intelligence as statistical inference." Glass Bead, vol. 1, no.1, pp. 25-36, 2017.

[18] L. B. Balzer and M. L. Petersen, "Invited commentary: machine learning in causal inference-how do I love thee? let me count the ways," American journal of epidemiology, vol. 190, no. 8, pp. 1483–1487, 2021.

[19] I. H. Sarker, "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions," SN computer science, vol. 2, no. 6, pp. 420–420, 2021.

[20] G. Jing, P. Li, Z. Chen and J. Zhang, "A survey on deep learning for multimodal data fusion," Neural Computation, vol. 32, no. 5, pp. 829-864, 2020.

[21] G. R. T. de Lima and G. B. Scofield, "Feasibility study on operational use of neural networks in a flash flood early warning system," Revista Brasileira de Recursos hídricos, vol. 26, no. 2, pp. 1-10, 2021.

[22] J. M. Silva, A. Figueiredo, J. Cunha, J.E. Dias, S. Silva et al., "Using rapid chlorophyll fluorescence transients to classify vitis genotypes," Plants (Basel), vol. 9, no. 2, pp. 174-189, 2020.

[23] M. Mostavi, Y.-C. Chiu, Y. Huang and Y. Chen, "Convolutional neural network models for cancer type prediction based on gene expression," BMC medical genomics, vol. 13, no.5, pp. 1–13, 2020.

[24] S.D. O'Donovan, K. essens, D. Lopatta, F. Wimmenauer, A. Lukas et al., "Use of deep learning methods to translate drug-induced gene expression changes from rat to human primary hepatocytes," Plos One, vol. 15, no. 8, pp. e0236392, 2020.

[25] H. Lahmer, A. E. Oueslati and Z. Lachiri, "Classification of DNA microarrays using deep learning to identify cell cycle regulated genes," 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), pp. 1–5, 2020.

[26] B. He, L. Bergenstrahle, L. Stenbeck, A. Abid, A. Andersson, et al., "Integrating spatial gene expression and breast tumour morphology via deep learning," Nature biomedical engineering, vol. 4, no. 8, pp. 827–834, 2020.

[27] O. Ahmed and A. Brifcani, "Gene expression classification based on deep learning," 4th Scientific International Conference Najaf (SICN), Najaf, Iraq, pp. 145–149, 2019

[28] K.W. Jordan, S. Wang, F. He, S. Chao, Y. Lun et al., "The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping," The Plant journal: for cell and molecular biology, vol. 95, no. 6, pp. 1039–1054, 2018.

[29] W. Hamdy, A. Darwish and A. E. Hassanien, "Artificial intelligence strategy in the age of covid-19: opportunities and challenges," Digital Transformation and Emerging Technologies for Fighting COVID-19 Pandemic: Innovative Approaches, Cham: Springer International Publishing, pp. 81–93, 2021.

[30] Y.C. Chiu, H. Chen, T. Zhang, S. Zhang, A. Gorthi et al., "Predicting drug response of tumors from integrated genomic

profiles by deep neural networks," BMC medical genomics, vol. 12, no. 1, pp. 18–155, 2019.

[31] H. Wang, R. Liu, P. Schyman and A. Wallqvist, "Deep neural network models for predicting chemically induced liver toxicity endpoints from transcriptomic responses," Frontiers in pharmacology, vol. 10, no. 4, pp. 42–42, 2019.

[32] A.H. El.Bassiouny, M. Abouhawwash and H.S. Shahen, "New generalized extreme value distribution and its bivariate extension," International Journal of Computer Applications, vol. 173, no. 3, pp. 1-10, 2017.

[33] A.H. El.Bassiouny, M. Abouhawwash and H.S. Shahen, "Inverted exponentiated gamma and its bivariate extension," International Journal of Computer Application, vol. 3, no. 8, pp. 13-39, 2018.

[34] A.H. El.Bassiouny, H.S. Shahen and M. Abouhawwash, "A new bivariate modified weibull distribution and its extended distribution," Journal of Statistics Applications & Probability, vol. 7, no.2, pp. 217-231, 2018.

[35] M. Abouhawwash and M.A. Jameel, "KKT proximity measure versus augmented achievement scalarization function," International Journal of Computer Applications, vol. 182, no. 24, pp. 1-7, 2018.

[36] H.S. Shahen, A.H. El-Bassiouny and M. Abouhawwash, "Bivariate exponentiated modified weibull distribution," Journal of Statistics Applications & Probability, vol. 8, no. 1, pp. 27-39, 2019.

[37] M. Abouhawwash and M.A. Jameel, "Evolutionary multi-objective optimization using benson's karush-kuhn-tucker proximity measure," International Conference on Evolutionary Multi-Criterion Optimization, East Lansing, Michigan, USA, Springer, pp. 27-38, 2019.

[38] M. Abouhawwash, M.A. Jameel and K. Deb, "A smooth proximity measure for optimality in multi-objective optimization using benson's method," Computers & Operations Research, vol. 117, no. 2, pp. 104900, 2020.

[39] M. Abouhawwash, K. Deb and A. Alessio, "Exploration of multi-objective optimization with genetic algorithms for PET image reconstruction," Journal of Nuclear Medicine, vol. 61, no. 1, pp. 572-572, 2020.

[40] S. Ibrahim, H. Alhumyani, M. Masud, S.S. Alshamrani, O. Cheikhrouhouv et al., "Framework for efficient medical image encryption using dynamic S-boxes and chaotic maps," IEEE Access, vol. 8, no. 13, pp. 160433-160449, 2020.

[41] M. Rawashdeh, M. Zamil, S. M. Samarah, M. Obaidat and M. Masud, "IOT-based service migration for connected communities," Computers & Electrical Engineering, vol. 96, no. 2, pp. 1-10, 2021.

[42] M. AbdelBasset, R. Mohamed, M. Elhoseny, M. Abouhawwash, Y. Nam et al., "Efficient MCDM model for evaluating the performance of commercial banks: A case study," Computers, Materials & Continua, vol. 67, no. 3, pp. 2729-2746, 2021.

[43] B. Gomathi, S. Balaji, V.K. Kumar, M. Abouhawwash, S. Aljahdaliet al., "Multi-objective optimization of energy aware virtual machine placement in cloud data center," Intelligent Automation & Soft Computing, vol.33, no.3, pp. 1771-1785, 2022.

[44] M. Kumar, K. Venkatachalam, M. Masud and M. Abouhawwash, "Novel dynamic scaling algorithm for energy efficient cloud computing, " Intelligent Automation & Soft Computing, vol.33, no.3, pp. 1547-1559, 2022.

[45] R.S. Ram, K. Venkatachalam, M. Masud and M. Abouhawwash, "Air pollution prediction using dual graph convolution LSTM technique," Intelligent Automation & Soft Computing, vol.33, no.3, pp. 1639-1652, 2022.

[46] A.J. Basha, N. Rajkumar, M.A. AlZain, M. Masud and M. Abouhawwash, "Fog based self-sovereign identity with RSA in securing IoMT data," Intelligent Automation & Soft Computing, vol. 34, no. 3, pp.1693-1706, 2022.

[47] G. Ravikumar, K. Venkatachalam, M.A. AlZain, M. Masud and M. Abouhawwash, "Neural cryptography with fog computing network for health monitoring using IoMT," Computer Systems Science and Engineering, vol. 44, no.1, pp.945-959, 2023.

[48] R. Rajdevi, K. Venkatachalam, M. Masud, M.A. AlZain and M. Abouhawwash, "Proof of activity protocol for IoMT data security," Computer Systems Science and Engineering, vol. 44, no. 1, pp.339-350, 2023.

[49] G. Ravikumar, K. Venkatachalam, M. Masud and M. Abouhawwash, "Cost efficient scheduling using smart contract cognizant ethereum for IoMT," Intelligent Automation & Soft Computing, vol.33, no.2, pp. 865-877, 2022.

[50] N. Mittal, H. Singh, V. Mittal, S. Mahajan, A.K. Pandit et al., "Optimization of cognitive radio system using self-learning salp swarm algorithm," Computers, Materials & Continua, vol.70, no.2, pp.3821-3835, 2022.

[51] K.S. Bhuvaneshwari, J. Uma, K. Venkatachalam, M. Masud, M. Abouhawwash et al., "Gaussian support vector machine algorithm based air pollution prediction," Computers, Materials & Continua, vol.71, no.1, pp.683-695, 2022.

# Feature Fusion Based Deep Transfer Learning Based Human Gait Classification Model

C.S.S. Anupama[1], Rafina Zakieva[2], Afanasiy Sergin[3], E Laxmi Lydia[4], Seifedine Kadry[5,6,7], Chomyong Kim[8], and Yunyoung Nam[8,*]

[1]*Department of Electronics and Instrumentation Engineering, V.R.Siddhartha Engineering College, Vijayawada 520007, India*
[2]*Candidate of Pedagogical Sciences, Department of Industrial electronics and lighting engineering, Kazan State Power Engineering University, Kazan, 420066, Russia*
[3]*Department of Theories and Principles of Physical Education and Life Safety, North-Eastern Federal University named after M.K. Ammosov, Yakutsk, 677000, Russia*
[4]*Department of Computer Science and Engineering, GMR Institute of Technology, Andhra Pradesh, Rajam, India*
[5]*Department of Applied Data Science, Noroff University College, Kristiansand, Norway*
[6]*Artificial Intelligence Research Center (AIRC), Ajman University, Ajman, 346, United Arab Emirates*
[7]*Department of Electrical and Computer Engineering, Lebanese American University, Byblos, Lebanon*
[8]*Department of ICT Convergence, Soonchunhyang University, Asan, 31538, Korea*

**Contact: Yunyoung Nam. Email: ynam@sch.ac.kr*

*Abstract*—**Gait is a biological typical that defines the method by that people walk. Walking is the most significant performance which keeps our day-to-day life and physical condition. Surface electromyography (sEMG) is a weak bioelectric signal that portrays the functional state between the human muscles and nervous system to any extent. Gait classifiers dependent upon sEMG signals are extremely utilized in analysing muscle diseases and as a guide path for recovery treatment. Several approaches are established in the works for gait recognition utilizing conventional and deep learning (DL) approaches. This study designs an Enhanced Artificial Algae Algorithm with Hybrid Deep Learning based Human Gait Classification (EAAA-HDLGR) technique on sEMG signals. The EAAA-HDLGR technique extracts the time domain (TD) and frequency domain (FD) features from the sEMG signals and is fused. In addition, the EAAA-HDLGR technique exploits the hybrid deep learning (HDL) model for gait recognition. At last, an EAAA-based hyperparameter optimizer is applied for the HDL model, which is mainly derived from the quasi-oppositional based learning (QOBL) concept. A brief classifier outcome of the EAAA-HDLGR technique is examined under diverse aspects, and the results indicate improving the EAAA-HDLGR technique. The results imply that the EAAA-HDLGR technique accomplishes improved results with the inclusion of EAAA on gait recognition.**

## I. INTRODUCTION

Human gait, how an individual walks, is personally distinctive because of its physical properties difference betwixt individuals and might be employed as a novel biometric for the authentication and identification of a person [1]. In comparison to other biometrics, namely iris, face, and fingerprint, human gait recognition (HGR) has the advantage of non-invasion, non-cooperation (without any cooperation or interaction with the subject), hard to disguise and long distance, making it very attractive as a means of detection and demonstrates huge potential in the applications of surveillance and security [2, 3]. Many sensing modalities involving wearable devices, vision, and foot pressure were used for capturing gait data [4].

Traditional HGR techniques include data preprocessing and features extracted in a handcrafted manner for additional identification [5], frequently suffer from various challenges and constraints enforced by the difficulty of the tasks, namely occlusions, viewing angle, locating the body segments, shadows, large intra-class variations, etc. [6, 7]. Emerging trends in machine learning (ML), called deep learning (DL), have become apparent in the past few years as a ground-breaking tool for handling topics in computer vision, speech, sound, and image processing, tremendously outperforming virtually any baseline established previously [8]. The new model exempts the requirement of manually extracting representative features from the expert and provides primary outcomes based on HGR, surpassing the present difficulties and opening room for additional investigation [9, 10]. The manually extracted feature was affected by the smartphone's position while gathering the information [11]. Consequently, some typical statistical manually extracted feature was collected from raw smartphone sensor information. Afterwards, the extraction of handcrafted features, the shallow ML classifier, was used to identify many physical activities of the human. Therefore, shallow ML algorithm relies on handcrafted feature [12, 13]. The DL algorithm is more advanced than the shallow ML algorithm since the DL algorithm automatically learns useful features from the raw sensory information without human intervention and identifies the human's physical activities [14]. The shallow ML algorithm with DL algorithms and handcrafted extracting features with automatically learned features achieved better outcomes in carrying out the smartphone-based HGR model. So, it is apparent that integrating manually extracted

features with automatically learned features in the DL algorithm might enhance the potential of the smartphone-based HGR paradigm [15].

Khan et al. [16] developed a lightweight DL algorithm for the HGR method. The presented algorithm involves sequential step–pretrained deep model selection of feature classification. Initially, two lightweight pretrained methods are considered and finetuned with respect to other layers and freeze some middle layers. Thereafter, the model was trained by means of deep transfer learning (DTL) algorithm, and the feature was engineered on average pooling and fully connected layers. The fusion is carried out by means of discriminative correlation analysis, enhanced by the moth-flame optimization (MFO) technique. Liang et al. [17] examine the effect of every layer on parallel infrastructured CNN. In particular, we slowly freeze the parameter of GaitSet from high to low layers and see the performance of finetuning. Furthermore, the rise of the frozen layer has negative consequences on the performance; it could reach the maximal efficacy with a single convolution layer unfrozen.

The author in [18] proposed a novel architecture for HGR using DL and better feature selection. During the augmentation phase, three flip operations were performed. During feature extraction phase, two pre-trained models have taken place, NASNet Mobile and Inception-ResNetV2. These two models were trained and fine-tuned by means of the TL algorithm on the CASIA B gait data. The feature of the selected deep model was enhanced by means of an adapted three-step whale-optimized algorithm, and the better feature was selected. Hashem et al. [19] proposed an accurate and advanced end-user software system that is capable of identifying individuals in video based on the gait signature for the purpose of hospital security. TL algorithm based on pretrained CNN was used and capable of extracting deep feature vectors and categorizing people directly rather than a typical representation that includes hand-crafted feature engineering and computing the binary silhouettes.

In [20], proposed a novel fully automatic technique for HGR under different view angles using the DL technique. Four major phases are included: recognition using supervised learning methods, pre-processing of the original video frame, which exploits pre-trained Densenet-201 CNN model for extracting features, and decrease of further features in extraction vector based on a hybrid selection technique. Sharif et al. [21] suggest a method that efficiently deals with the problem related to walking styles and viewing angle shifts in real time. The subsequent steps are included: (a) feature selection based on the kurtosis-controlled entropy (KcE) method, then a correlation-based feature fusion phase, (b) realtime video capture, and (c) extraction feature utilizing TL on the ResNet101 deep model. Then, the most discriminative feature was categorized by means of the innovative ML classifier.

This study designs an Enhanced Artificial Algae Algorithm with Hybrid Deep Learning based Human Gait Classification (EAAA-HDLGR) technique on sEMG signals. The EAAA-HDLGR technique initially extracts the time domain (TD), and frequency domain (FD) features from the sEMG signals and is fused together. In addition, the EAAA-HDLGR technique exploits the hybrid deep learning (HDL) model for gait recognition. At last, an EAAA-based hyperparameter optimizer is applied for the HDL model, mainly derived from the quasi-oppositional based learning (QOBL) concept. A brief classifier outcome of the EAAA-HDLGR technique is examined under diverse aspects.

## II. The Proposed Model

In this study, we have derived a new EAAA-HDLGR technique for gait recognition using sEMG signals. Primarily, the EAAA-HDLGR technique derived the TD as well as FD features from the sEMG signals, which are then fused together. In addition, the EAAA-HDLGR technique exploited the HDL model for gait recognition. At last, an EAAA-based hyperparameter optimizer is applied for the HDL model, mainly derived from the QOBL concept. Fig. 1 depicts the workflow of the EAAA-HDLGR algorithm.

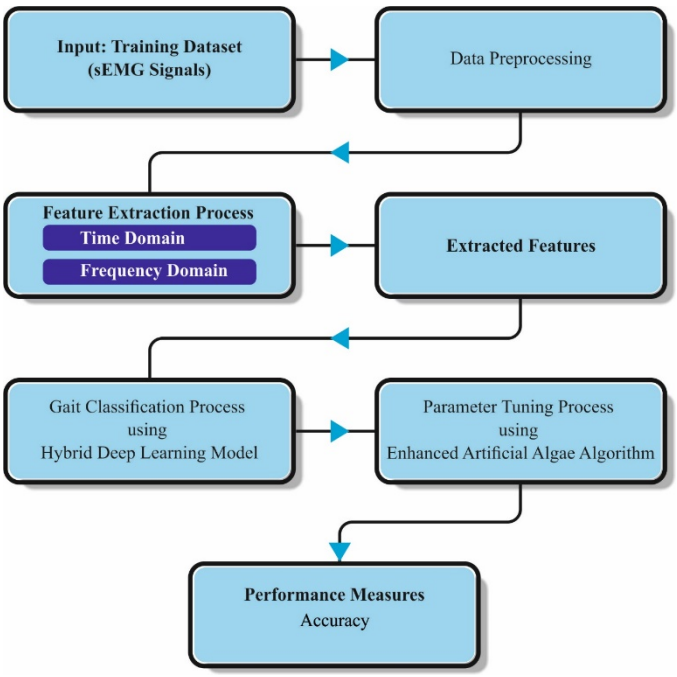



Fig. 1. Workflow of EAAA-HDLGR approach

### A. Feature Extraction Process

Afterwards, de - noising, the TD and FD features of all the channels of the EMG signal can be extracted [22]. During this case, the 3 representative time domain features comprising variance (VAR), zero crossing points (ZC), and mean absolute value (MAV) can be utilized as frequency domain features. MAV gets the benefit of properties in which sEMG signals are huge amplitude fluctuations from the time domain that are linearly compared with muscle activation level. The maximum value of MAV is the superior activation level of muscles.

$$MAV = \frac{1}{N}\sum_{k=1}^{N} |x_k| \quad (1)$$

Whereas, $x_k(k = 1,2,\dots,N)$ implies the sEMG time series with a window length of $N$. VAR is the size of signal powers of sEMG signals and is formulated as:

$$VAR = \frac{1}{N-1}\sum_{k=1}^{N} x_k^2 \quad (2)$$

$ZC$ implies the count of times the sEMG waveform passes with the zero point to avoid signal cross-calculating due to low-level noise. It can be a mathematical process as follows:

$$ZC = \sum_{k=1}^{N} sgn(-x_k x_{k+1}) \quad (3)$$

$$\text{whereas, } sgn(x) = \begin{cases} 1 \; x > 0 \\ 0 \; otherwise \end{cases}$$

It can choose 2 representative frequency domain features like average power frequency $f_{mean}$ and median frequency $f_{mf}$ determined as:

$$f_{mean} = \frac{\int_0^{+\infty} f\,P(f)df}{\int_0^{+\infty} P\,(f)df} \quad (4)$$

$$\int_0^{f_{mf}} P\,(f)df = \int_{f_{mf}}^{+\infty} P\,(f)df = \frac{1}{2}\int_0^{+\infty} P\,(f)df \quad (5)$$

Whereas $P(f)$ implies the power spectral density of sEMG signals, and $f$ denotes the frequency.

### *B. Gait Classification using HDL Model*

During this study, the HDL technique was employed for the gait classifier. It comprises CNN with LSTM and Bi-LSTM technique [23]. CNN is increasingly popular under the domains of DL and is contained in convolution (Conv) and pooling layers. The Conv layer function is to remove useful features in the input database. Its interior includes one or more Conv kernels. The Conv layer was executed to remove effectual features by sliding the Conv kernel on the feature. Afterwards, enhance a Max-Pooling layer and then the Conv layer; the Max-Pooling layer maintains the strong features and extracts the weaker feature to prevent over-fitting and decrease the complexity. LSTM avoids gradient vanishing and explosion from the trained process of RNN, whereas input, output, and forget gates can be projected. The 3 gate functions effectively solve the trained problem of RNN. The input gate determines the data that existing time carries to future time. The forget gate identifies the data count of preceding times which is preserved in the current period. The resultant gate chooses the resultant of existing to future states. The succeeding equation illustrates the equation of distinct cells from LSTM:

$$i_z = \sigma(W_{xi}x_z + W_{hi}h_{z-1} + b_i) \quad (6)$$

$$f_z = \sigma\left(W_{xf}x_z + W_{hf}h_{z-1} + b_f\right) \quad (7)$$

$$o_z = \sigma(W_{xo}x_z + W_{ho}h_{z-1} + b_o) \quad (8)$$

$$\tilde{C}_z = \tanh\,(W_{xc}x_z + W_{hc}h_{z-1} + b_c) \quad (9)$$

$$c_z = f_z e c_{z-1} + i_z e \tilde{C}_z) \quad (10)$$

$$h_z = o_z e \tanh\,(c_z) \quad (11)$$

At this point, $W_{hf}, W_{hc}, W_{ho}, W_{hi}$ implies the weighted in hidden layer to forget, memory cell, output, and input gates. $W_{xf}, W_{xo}, W_{xi}$, and $W_{xc}$ signify the weight of forgetting gate, output gate, input gate, and memory cell. $f_z, i_z, 0_z$ illustrate the $z^{th}$ forget, input, and output gates. $b_c\ b_o, b_f, b_i$ stands for the bias values of a memory cell, output gate, forget gate, and input gate. Bi-LSTM is a group of forward and backward LSTM that could provide relevant data in forward and backward ways and concatenate the prediction. LSTM is suitable the time-related data in one way. BiLSTM improves the opposite way of LSTM in such a method that BiLSTM captures the pattern, which LSTM ignores. LSTM-BiLSTM classifier the gait based on extracting feature.

- CNN was projected with three 1-D Conv layers, and the count of Conv kernels is set as 16, 32, 64, correspondingly. The Conv kernel size is set as 2, and the striding phase has set as. The effectual feature was extracted with stride from the Conv kernel. Then, add the Max-Pooling layer for all the Conv layers; afterwards, the pooling window size was set as 2, and the stride stage was set as one. The Max-Pooling layer decreases the feature effort and avoids overfitting.
- By assigning weight to features using the attention process, the attention block enhances the outcome of time series features, restrains the interference of insignificant features, and effectually explains this problem that process isn't judge the effect of vital of various time series features.
- The extraction feature assumed that input of 1 Bi-LSTM layer and 2 LSTM layers achieved the classifier outcome.

### *C. Parameter Optimization using EAAA Model*

To improve the recognition rate, EAAA based hyperparameter optimizer is applied to the HDL model. Artificial algae are generally known to describe the features of algae and demonstrate that they can be responsive to solutions from the problem spaces [24]. In relation to a real one, artificial algae demonstrate that if it is implemented in environment by moving to light source for photosynthesizing by spiral swimming, it is able to switch superior species and eventually multiply with mitotic division. Thus, this procedure contains 3 important processes Helical Motion, Evolutionary Process, and Adaptation. The term algae colony signifies the collection of algae cells that lives together. Algae colony and population can demonstrate in the subsequent formula.

$$Population = \begin{bmatrix} \chi_1^1 & \cdots & \chi_1^D \\ \vdots & \ddots & \vdots \\ \chi_N^1 & \cdots & \chi_N^D \end{bmatrix} \quad (12)$$

$$Algae\ colony = [\chi_i^1, \chi_i^2\ \chi_i^D] \quad (13)$$

The algae colony work as separate cell and permits together, and the cell from the colony can pass away an unfavorable level of life. The colony present at the optimal point is named a better colony, and it includes better algae cells. The development kinetics of algae colonies calculated by Monod approach as demonstrated under.

$$\mu = \frac{\mu_{\max} S}{K_s + S} \tag{14}$$

In Eq. (14), $\mu$ signifies the rate of growth, $\mu$ max implies the maximal rate of growth, $S$ represents the nutrient concentration, and $K$ is divided as to algal colony. In Eq. (17), $\mu$max value is set to one (dependent upon mass conservation model, an entire count converted as to biomass can correspond to the count of substrate expended to all the time units. The size of $i^{th}$ algae colonies at time $t+1$ from the Monod equation is depicted under:

$$G_i^{t+1} = \mu_i^t G_i^t \; i = 1,2,N \tag{15}$$

In Eq. (15), $G_i^t$ denotes the size of $i^{th}$ algae colonies at $t$ time, $N$ represents the amount of algae colonies. The algae colony offers optimum solution that develops higher as a sign of greater count of nutrients is obtained. In all the algae cells, the minimal algae colony dies from the evolution procedure, and the algae cell of maximum algae colony reproduces. Such procedure is complete in the subsequent formulas:

$$biggest^t = \max G_i^t i = 1,2,\cdots N \tag{16}$$

$$mallest^t = \min G_i^t i = 1,2,\cdots N \tag{17}$$

$$smallesl_m^t = biggest_m^t m = 1,2,\cdots D\; D \tag{18}$$

In the above formula, D signifies that there is a problem feature, whereas the maximum describes the largest colony of algae, and the minimal symbolizes the lower colony of algae. A primary hunger value is 0 for every artificial alga, and Adaptation process stops by an alteration in the hunger level.

$$starving^t = \max A_i^t i = 1,2,\cdots N \tag{19}$$

$$starving^{t+1} = slarving^t + (bigges^t - starving^t) rand \tag{20}$$

During this equation, the value $A$ signifies the hunger value of $i^{th}$ algae colonies at $t$ time, and starving$^t$ describes the algae colonies with maximal angle value at $t$ time. In Eq. (20), the variation parameter states the adaptation scheme and is utilized at $t$ time. Generally, the variation parameter value lies in the interval of 0 and 1. In AAA, the movement of algae cells are spiral, the viscous drag is realized as shear force commensurate to algae cell size, gravity restricting movement is presented as zero, and viscous drag demonstrates as zero. At this time, $\tau(X)$ implies the friction surface, the friction surface is the surface area, and Algae colony has spherical from the shape demonstrated in the under equation. Fig. 2 depicts the flowchart of AAA.

$$\tau(x_i) = 2\pi r^2 \tag{21}$$

$$\tau(x_i) = 2\pi \left( 3\sqrt{\frac{3G_i}{4\pi}} \right)^2 \tag{22}$$

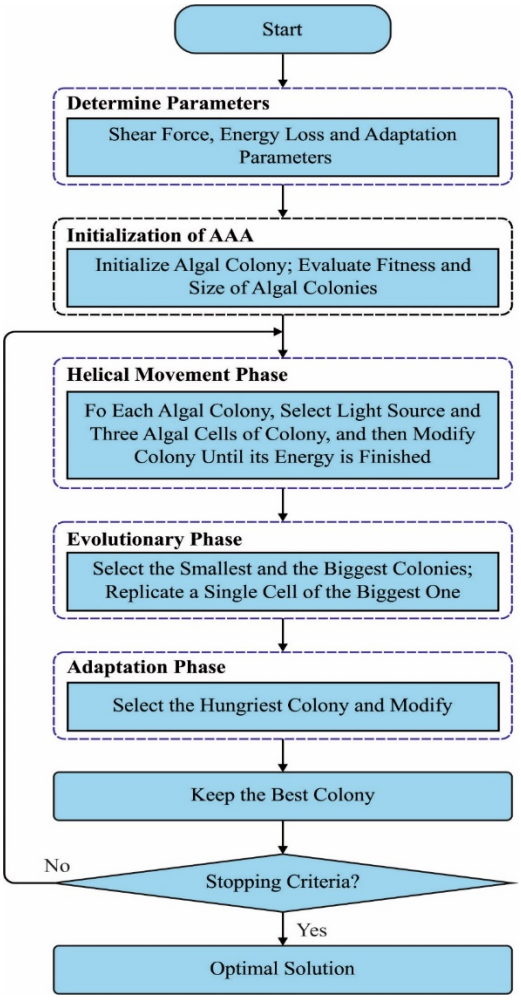



Fig. 2. Flowchart of AAA

The distance to light source and friction surface is to control the step size of motions.

$$x_{im}^{t+1} = x_{im}^t + (x_{jm}^t - x_{im}^t)(\Delta - \tau^t(X))p \tag{23}$$

$$x_{ik}^{t+1} = x_{ik}^t + \left(x_{jk}^t - x_{ik}^t\right)\left(\Delta - \tau^t(x_i)\right)\cos\alpha \tag{24}$$

$$x_{i1}^{t+1} = x_{i1}^t + \left(x_{j1}^t - x_{i1}^t\right)\left(\Delta - \tau^t(x_i)\right)\sin\beta \tag{25}$$

For the helical rotation of algae cells, ($x_{im}^t, x_{ik}^t$, and $x_{i1}^t$) are stated as the co-ordinate of the algae cell ($x, y$, and z) at $t$ time. $\alpha$ and $\beta \in [0,2]; p \in [-1,1]; \Delta$ the force to procedure; $t$ (xi) signifies the friction surface area of $i^{th}$ algae cells. The AAO technique arbitrarily creates the primary solution in the search range. Once the solution is superior to the global better solution, afterward, the convergence population rate can be slower and simply decrease as to the local better solution. Tizhoosh projected oppositional-based learning (OBL) technique for preventing these local better problems [25].

Also, the QOAAO technique resulted by utilizing quasi-opposition based population initialization. At this point, the random solution was higher than global better solution if related to their opposite solution. So, the N better individual was

chosen in the distinct population containing $N$ arbitrary individuals and the opposite solution of individuals. Let $X = (X_1, X_2, \dots, X_D)$ refers to the point in $d$ dimension space, and its opposite point signifies $X^{OBL} = (X_1^{OBL}, X_2^{OBL}, \dots, X_D^{OBL})$ which is calculated by the subsequent formula; and its quasi-opposite point $X^{QOBL} = \left(X_1^{QOBL}, X_2^{QOBL}, \dots, X_D^{QOBL}\right)$ was measured as:

$$X_d^{OBL} = lb_d + ub_d - X_d, \tag{26}$$

$$X_d^{QOBL} = \begin{cases} \frac{lb_d+ub_d}{2} + rand(0,1) \times (X_d^{OBL} - \frac{lb_d+ub_d}{2}) X_d < \frac{lb_d+ub_d}{2} \\ X_d^{OBL} + rand(0,1) \times (\frac{lb_d+ub_d}{2} - X_d^{OBL}) X_d \geq \frac{lb_d+ub_d}{2} \end{cases}, \tag{27}$$

In the formula, the $d$ dimension vector of $X$ was signified as $X_d$, $X_d^{OBL}$, and $X_d^{QOBL}$ denotes the $d$ dimensional opposite point $X^{OBL}$, and quasi-opposite point $X^{QOBL}$ correspondingly. Also, $lb_d$ and $ub_d$ illustrates the $d$ dimensional of lower as well as upper limits of problems, correspondingly.

The fitness chosen is a vital feature in the EAAA system. The solution encoder was exploited to assess the aptitude (goodness) of candidate solutions. At this point, the accuracy value is main form employed to design a fitness function.

$$Fitness = \max(P) \tag{28}$$

$$P = \frac{TP}{TP+FP} \tag{29}$$

From the expression, TP represents the true positive, and FP denotes the false positive value.

## III. Results and Discussion

In this section, the gait classification results of the EAAA-HDLGR approach are investigated in detail. Tab. 1 and Fig. 3 demonstrate the overall gait classification outcomes of the EAAA-HDLGR model with other ML models on TD features [22]. The experimental results inferred that the EAAA-HDLGR model exhibits effectual outcomes. For instance, with pre-stance stage, the EAAA-HDLGR model has highlighted increasing accu_y of 96.07% while the SVM, ELM, LSTM, DBN, and SSA-DBN models have attained reducing accu_y of 95%, 96.07%, 93.49%, 93.63%, and 96.99% respectively. Meanwhile, with Terminal-Stance stage, the EAAA-HDLGR approach has emphasized increasing accu_y of 98.95% while the SVM, ELM, LSTM, DBN, and SSA-DBN models have attained decreasing accu_y of 92.01%, 95.50%, 97.33%, 94.73%, and 95.78% correspondingly. Finally, with Terminal-Swing stage, the EAAA-HDLGR technique has demonstrated higher accu_y of 98.03% while the SVM, ELM, LSTM, DBN, and SSA-DBN methods have attained lesser accu_y of 96.92%, 94.79%, 92.76%, 96.65% and 94.84% correspondingly.

**TABLE I. GAIT CLASSIFIER OUTCOME OF EAAA-HDLGR APPROACH WITH OTHER ML TECHNIQUES ON TD FEATURES**

| Accuracy (%) - TD Features | | | | | | |
|---|---|---|---|---|---|---|
| **Class** | **SVM** | **ELM** | **LSTM** | **DBN** | **SSA-DBN** | **EAAA-HDLGR** |
| Pre-Stance | 95.00 | 96.07 | 93.49 | 93.63 | 96.99 | 96.07 |
| Mid-Stance | 95.15 | 94.45 | 95.89 | 92.36 | 96.49 | 95.56 |
| Terminal-Stance | 92.01 | 95.50 | 97.33 | 94.73 | 95.78 | 98.95 |
| Pre-Swing | 92.17 | 92.77 | 93.89 | 92.65 | 92.03 | 95.78 |
| Terminal-Swing | 96.92 | 94.79 | 92.76 | 96.65 | 94.84 | 98.03 |
| **Average** | **94.25** | **94.72** | **94.67** | **94.00** | **95.23** | **96.88** |

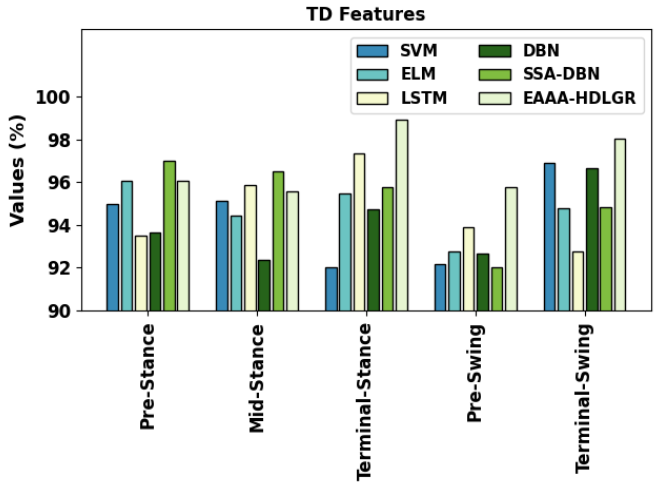



Fig. 3. Gait classifier outcome of EAAA-HDLGR approach on TD features

## IV. Conclusion

In this study, we have derived a new EAAA-HDLGR technique for gait recognition using sEMG signals. Primarily, the EAAA-HDLGR technique derived the TD as well as FD features from the sEMG signals which are then fused together. In addition, the EAAA-HDLGR technique exploited the HDL model for gait recognition. At last, EAAA based hyperparameter optimizer is applied for the HDL model, which is mainly derived by the use of the QOBL concept. A brief classifier outcome of the EAAA-HDLGR technique is examined under diverse aspects and the results indicated the betterment of the EAAA-HDLGR technique. The results imply that the EAAA-HDLGR technique accomplishes improved results with the inclusion of EAAA on gait recognition. In future, feature reduction and feature selection processes can be combined to boost the recognition rate of the EAAA-HDLGR technique.

## Acknowledgment

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number :

HI21C1831) and the Soonchunhyang University Research Fund.

## REFERENCES


[1] Z. Ni and B. Huang, "Human identification based on natural gait micro-Doppler signatures using deep transfer learning," IET Radar, Sonar & Navigation, vol. 14, no. 10, pp. 1640-1646, 2020.

[2] X. Bai, Y. Hui, L. Wang and F. Zhou, "Radar-based human gait recognition using dual-channel deep convolutional neural network," IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 12, pp. 9767-9778, 2019.

[3] M. Sharif, M. Attique, M.Z. Tahir, M. Yasmim, T. Saba et al. "A machine learning method with threshold based parallel feature fusion and feature selection for automated gait recognition," Journal of Organizational and End User Computing, vol. 32, no. 2, pp. 67-92, 2020.

[4] H. Wu, X. Zhang, J. Wu and L. Yan, "Classification algorithm for human walking gait based on multi-sensor feature fusion," in International Conference on Mechatronics and Intelligent Robotics, Advances in Intelligent Systems and Computing book series, Springer, Cham, vol. 856, pp. 718–725, 2019.

[5] H. Arshad, M.A. Khan, M. Sharif, M. Yasmin and M.Y. Javed, "Multi-level features fusion and selection for human gait recognition: an optimized framework of Bayesian model and binomial distribution," International Journal of Machine Learning and Cybernetics, vol. 10, no. 12, pp. 3601-3618, 2019.

[6] N. Mansouri, M. A. Issa and Y. B. Jemaa, "Gait features fusion for efficient automatic age classification," IET Computer Vision, vol. 12, no. 1, pp. 69-75, 2018.

[7] D. Thakur and S. Biswas, "Feature fusion using deep learning for smartphone based human activity recognition," International Journal of Information Technology, vol. 13, no. 4, pp. 1615-1624, 2021.

[8] F.M. Castro, M.J. Marín-Jiménez, N. Guil and N. Pérez de la Blanca, "Multimodal feature fusion for CNN-based gait recognition: an empirical comparison," Neural Computing and Applications, vol. 32, no. 17, pp. 14173-14193, 2020.

[9] M.M. Hasan and H.A. Mustafa, "Multi-level feature fusion for robust pose-based gait recognition using RNN," International Journal of Computer Science and Information Security, vol. 18, no. 2, pp. 20-31, 2021.

[10] Y. Lang, Q. Wang, Y. Yang, C. Hou, D. Huang et al. "Unsupervised domain adaptation for micro-Doppler human motion classification via feature fusion," IEEE Geoscience and Remote Sensing Letters, vol. 16, no. 3, pp. 392-396, 2018.

[11] M. Li, S. Tian, L. Sun and X. Chen, "Gait analysis for post-stroke hemiparetic patient by multi-features fusion method," Sensors, vol. 19, no. 7, pp. 1737, 2019.

[12] A. Abdelbaky and S. Aly, "Two-stream spatiotemporal feature fusion for human action recognition," The Visual Computer, vol. 37, no. 7, pp. 1821-1835, 2021.

[13] F.F. Wahid, "Statistical features from frame aggregation and differences for human gait recognition," Multimedia Tools and Applications, vol. 80, no. 12, pp. 18345-18364, 2021.

[14] Q. Hong, Z. Wang, J. Chen and B. Huang, "Cross-view gait recognition based on feature fusion," in IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI), Washington, DC, USA, pp. 640-646, 2021

[15] K. Sugandhi, F.F. Wahid and G. Raju, "Inter frame statistical feature fusion for human gait recognition," in International Conference on Data Science and Communication (IconDSC), Bangalore, India, pp. 1-5, 2019

[16] M.A. Khan, H. Arshad, R. Damaševičius, A. Alqahtani, S. Alsubai et al. "Human gait analysis: A sequential framework of lightweight deep learning and improved moth-flame optimization algorithm," Computational Intelligence and Neuroscience, vol. 2022, pp. 1–13, 2022.

[17] Y. Liang, E.H.K. Yeung and Y. Hu, "Parallel CNN classification for human gait identification with optimal cross data-set transfer learning," in IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Hong Kong, China, pp. 1-6, 2021

[18] F. Saleem, M.A. Khan, M. Alhaisoni, U. Tariq, A. Armghan et al. "Human gait recognition: A single stream optimal deep learning features fusion," Sensors, vol. 21, no. 22, pp. 7584, 2021.

[19] L. Hashem, R. Al-Harakeh and A. Cherry, "Human gait identification system based on transfer learning," in 21st International Arab Conference on Information Technology (ACIT), Giza, Egypt, pp. 1-6, 2020.

[20] A. Mehmood, M.A. Khan, M. Sharif, S.A. Khan, M. Shaheen et al. "Prosperous human gait recognition: an end-to-end system based on pre-trained CNN features selection," Multimedia Tools and Applications, pp. 1-21, 2020, doi: 10.1007/s11042-020-08928-0

[21] M.I. Sharif, M.A. Khan, A. Alqahtani, M. Nazir, S. Alsubai et al. "Deep learning and kurtosis-controlled, entropy-based framework for human gait recognition using video sequences," Electronics, vol. 11, no. 3, pp. 334, 2022.

[22] J. He, F. Gao, J. Wang, Q. Wu, Q. Zhang et al. "A method combining multi-feature fusion and optimized deep belief network for emg-based human gait classification," Mathematics, vol. 10, no. 22, pp. 4387, 2022.

[23] R.J. Kavitha, C. Thiagarajan, P.I. Priya, A.V. Anand, E.A. Al-Ammar et al. "Improved harris hawks optimization with hybrid deep learning based heating and cooling load prediction on residential buildings," Chemosphere, vol. 309, pp. 136525, 2022.

[24] K.I. Anwer and S. Servi, "Clustering method based on artificial algae algorithm," in International Journal of Intelligent Systems and Applications in Engineering, vol. 9, no. 4, pp. 136-151, 2021.

[25] J. Xia, H. Zhang, R. Li, Z. Wang, Z. Cai et al. "Adaptive barebones salp swarm algorithm with quasi-oppositional learning for medical diagnosis systems: a comprehensive analysis," Journal of Bionic Engineering, vol. 19, no. 1, pp. 240-256, 2022.

# Fall Risk Assessment from Long-Term Gait Data of Inertial Sensor using Deep Learning

Kuntha. Pin[1, *], Jung-Yeon. Kim[2], Yunyoung. Nam[3]
[1]*Department of ICT Convergence, Soonchunhyang University, Asan, South Korea*
[2]*ICT Convergence Research Center, Soonchunhyang University, Asan, South Korea*
*Department of Computer Science and Engineering, Soonchunhyang University, Asan, South Korea*
*Contact: pin.kuntha145@gmail.com

*Abstract*— **Fall is a critical risk for the elderly and patients, especially patients with neurodegenerative disease. People with falls possibly have a probability of a weaker risk of death. Fall risk assessments have been proposed to eliminate and prevent falls using various methods, and accuracy needs to improve. Moreover, due to the development of technology, a wearable sensor like an inertial measurement (IMU) sensor can capture the movement's characteristics. Therefore, this study applies a deep learning-based computer intelligence method to assess the level of fall into low and high levels through analysis of long-term gait signals from IMU sensors. IMU sensors are placed on the ankle of patients to record the movements of feet movement. The analysis gait signal is divided into four steps. First, the gait signal is proposed by performing an imputation operation and outlier removal to remove the missing and outlier values from the signal. Then gait signal is hourly segmented, better representing the detail of gait data. Secondly, the gait phases, toe-off and heel-strike are detected from each segmented gait. Third, feature extraction is used to extract the gait parameters based on the detected gait phases and to concatenate the gait parameters of each piece. Finally, sequential gait parameters representing walking are utilized as guiding features for the long short-term (LSTM) model in classifying the levels of falling. Long-term recording data uses a lot of memory. Hence, this study tries to reduce using storage trough extracting time sequence gait parameters and details the effectiveness of features to obtain better performance. The better performance of fall-level detection can help clinically track and informing to prevent falls.**

## I. INTRODUCTION

The of falling of the risk assessment is the future prediction of falls and early threats to prevent falls. Most of falling in the elderly and patients, especially patients who have Parkinson's disease (PD) [1]. The level of risk indicates a significant health status for clinical treatment and providing effective advice to the patient. Fall risk is divided into two levels, low and high fall risk [2]. Low fall risk level refers to less probability of falling, and for high fall risk, the brain has less ability to control gait, which results in more falls. Developing technology like wearable devices attached to people's bodies for capturing characteristics of change during walking.

Inertial Measurement Unit (IMU) sensor is a wearable device widely used to track an object's motion by measuring the direction of movement according to time. In the medical field, it is attached to specific parts of the body to capture activities. For example, to capture people's walking, it is placed on various positions like lower back, thigh, ankle, instep, etc. [3-5]. However, the most accurate and best placement sensor is on the instep or ankle [6] because it can represent a change of gait.

The duration of using the IMU sensor to record gait data is essential in the fall risk assessment study. Short-time recording and restricted walking may be prone to errors when applied in real life. Long-term data recording over multiple days with unconstrained walking is valuable and possible in accessing fall risk. It is a kind of free life walking and represents the health condition in long time.

The gait signal is a continuous real-time signal; the gait cycle is presented in a period. Computer intelligence methods, machine learning and deep learning have been used to detect gait phases or predict the change between healthy and non-healthy people. Especially deep learning based on long-short-term memory (LSTM) has been popularly used to analyse continuous data for a long time. Remembering the previous sequence data input in predicting the next task is possible.

In this study, the fall risk assessment method is proposed to assess the fall risk into two levels, low and high fall risk, using long-term data and deep learning based on the LSTM model to analyze the falling level. First, IMU sensors are placed on the ankle of the patient's feet to record daily gait data. Next, long-term data is segmented hourly to extract specific features, which is better than extracting features from raw data. Then hourly-Segmented is passed to the next processing to access the gait and non-gait data, remove unnecessary data resulting in reduced computation time, and improve the accuracy of predicting in the next step. Pre-processing and gait phases are applied to filter out the noise and smooth and access gait events. The gait parameters of each segmentation are extracted on gait events and concatenated together to contract a sequence gait parameter. Finally, deep learning based on the LSTM model is used to analyze the sequence of gait parameters to assess the risk level into low and high risk of falls.

The contribution of this study is as follows:

- Detection of gait and non-gait signals from the long-term gait signal.
- Proposed an algorithm to detect gait phases, including toe-offs and heel strikes, from the gait signal.

- Extraction of gait parameters, such as stride time, swing time, and stance time, based on the detected gait phases.
- Utilization of the extracted features for classifying the level of fall risk.

## II. Literature Review

Fall risk assessments have been studied using various computer-intelligent methods to assess the fall risk level. Tunca et al. [7] used deep learning based on the LSTM model to classify low and high fall risk using unconstrained gait. The patients needed to wear IMU sensors placed on the instep of both feet and walk straight for 8 meters many times. Gait parameters and perform data augmentation with 10-stride sliding windows. The features were extracted and prepared into two algorithms for training machine learning and deep learning models. As a result, the LSTM model with sequential input of gait parameters performs with a balanced accuracy of 92.10%. For another study, Ullrich et al. [8] used various machine learning algorithms to classify low and high fall risk levels using long-term data recorded from an IMU sensor (placed in the instep of both feet). Before classification, initial contact, mid-stance, and final contact were detected, extracting various spatial gap parameters and aggregating features to train machine learning models. Support vector machine (SVM), Random Forest (RF), and Gradient boosting (GB) were trained with different numbers and types of features. Among those used models, the RF model achieved a higher performance with an accuracy of 74.0%. The development of previous methods shows that the results need to improve for greater accuracy and applicability in the real world. Therefore, this study aims to eliminate those lacking by applying the LSTM model to analyse a sequence of gait parameters from long-term gait data.

## III. Methodology

The proposed fall risk assessment method is broken down into steps, as shown in Fig.1. Long-term gait data is segmented in a short period. Then, before passing the signal to the next process, the gait and non-gait were detected. The next processing is pre-processing to remove non-values and smooth the signal. Next, gait phases like toe-off and heel-strike are detected. Next, the gait parameters are calculated according to the gait phase points. Then, the sequence of hourly gait parameters is concatenated to construct the sequential data. Finally, the LSTM model is applied to analyse the fall risk level. For further detail of each process in the below sections.

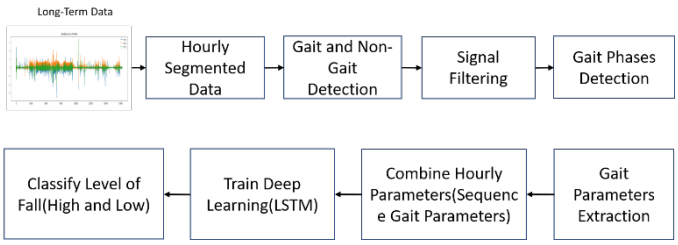



**Figure 1.** Flowchart of the proposed fall risk assessment method.

### A. Dataset

The dataset was collected from the Soonchunhyang Buchoen Hospital. The doctor labelled patient's data into low or high fall risk based on statistical records, patients with experience with falls were labelled as high fall risk and without experience were labelled as low fall risk level. The data was captured by IMU sensors (AX3) placed on the instep of both feet and represented the gait by three accelerations along with the x-axis, y-axis, and z-axis. Gait data was captured at a sampling frequency of 100. Y-axis refers to the vertical direction, and the x-axis is for the forward and backward direction.

### B. Signal Pre-processing

The input data is recorded multiple hours with a single label (the label could be high or low fall risk).

#### 1. Signal Segmentation

The multiple-hour data is segmented into hourly in order to start from recording time to stopping time per day. It is simply to visualize the signal in a short period and at a specific time for each patient. The total segmented data for each patient is given in Eq. (1).

$$S_p = [s_{t+1}, s_{t+2}, s_{t+3}, \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots s_{t+n}] \quad (1)$$

Where p denotes the patient's identity, s represents the one-hour segmented data at a specific hour, t ($0 \leqslant t + n \leqslant 23$) denotes the start time, and n represents the number of hours of recording.

#### 2. Gait and Non-Gait Detection

A single data contains multiple actions because patients are freely walking without any limitations. The signal from those actions can be categorized into two main categories, gait and non-gait. The gait signal refers to a signal that contains any pattern when walking or running, such as toe-off, heel-strike, etc. Non-gait signal is a signal without gait phases, for example, when sleeping, standing, or any actions. Therefore, we proposed our algorithm to detect gait and non-gait signals. The detection algorithm is an examination of the signal at y-acceleration. Distance between two higher peaks, and distance between two lower peaks. The highest peak is detected by using the SciPy Python library with setting parameters: the distance from each peak is 25 samples, the minimum height is 0.5 and with 5 samples. Then, according to higher peaks, the lower peaks are detected using the NumPy library; the local minima point between two higher peaks is a lower peak. The equations to find higher peaks and lower peaks are provided in Eq. (2) and (3), respectively. When the distance from one lower peak to the next lower peak is less than 100 samples, it is labelled as a gait signal (represented by the number 1). Conversely, if the distance from a higher peak to the next higher peak is less than 100 samples, it is also labelled as a gait signal. In contrast, when the distance exceeds 100 samples in either case, the signal is labelled as a non-gait signal (represented by the number 0). Detected results according to peak presentation and peak distance are shown in Fig 2. Gait signals with peaks of less than 6 are relabelled as non-gait signals. The final result of detecting gait and non-gait signals is shown in Fig. 3.

hp-indices = Y (h, th, dst, p, w, r, plt_size), (2)

Where h denotes the height of peaks, th represents the threshold of peaks, dst denotes the distance from one peak to the next nearby peak, p represents the prominence, w denotes the width of peaks, r represents the peak width, and plt_size indicates the number of samples on the flat top of the peaks.

$$\text{lp} = \text{np.argmin}([s_{hp-indices[i]}, s_{hp-indices[i+1]}]) \quad (3)$$

Where $s$ stands for one-hour segmented data at a specific hour and at the higher peak index i, i is the iteration index (i≤number peaks-1).

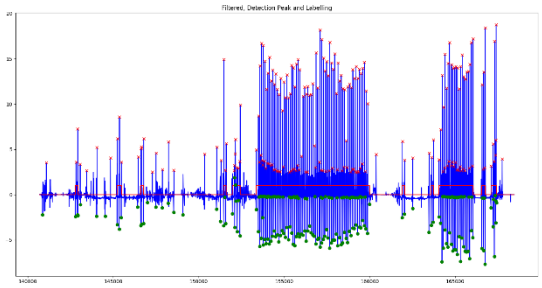


**Figure 2.** Gaits and non-gait labelling: Example of labelling based on peak distance for a high-risk signal of one hour.

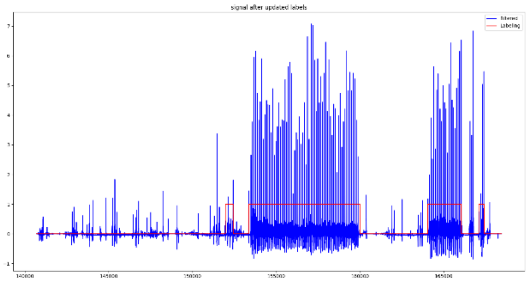


**Figure 3.** Gaits and non-gaits labelling: Example of final labeling for high-risk signal of one hour.

## 3. Signal Filtering

The recording of data from sensors may have some failures. The missing values cause difficulties in analysis or cannot make a computation [9]. Therefore, gait data is performed in imputation to remove the missing values and fill them with zero values instead. Moreover, the gait signal also contains noise at a higher frequency. In this study, the effective axis is the y-axis because it contains higher values than the other two. The gaussian filter is applied to remove noises and make data smoother. The gaussian function is given in eq. (4). Smooth signal by sigma of $5/\sqrt{8 \times \log(2)}$. The comparison gait signal before and after applying the Gaussian filter is shown in Fig. 2.

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{x^2}{2\sigma^2}} \quad (4)$$

Where $\sigma$ is the standard deviation of the distribution. x denotes one-dimensional data input.

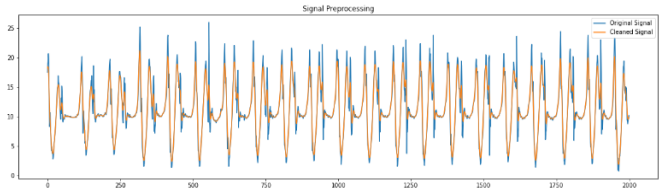


**Figure 4.** Example of a smoothed gait signal for segmented data over 20 seconds.

### C. Gait Phases Detection

Gait phases are an essential key in the analysis of the gait signal. To access the specific change of gait in time. Many studies detected gait phases first and then constructed the features representing the gait signal that varies in time and at each level of disease [10]. Toe-off and heel-strike are gait phases that are more commonly detected by previous studies because the signal change at this event changes more than another gait event, which is to be distinguished. Studying the y-axis shows that toe-off is the first local maxima and heel-strike is the second local maxima; both order each other for actual walking [11]. In this study, we used the *find_peaks* function of the Scipy Python library to detect local maxima. Then, we proposed our algorithm to classify toe-off and heel-strike from those local maxima indices.

The proposed algorithm for classifying two gait phases refers to a real examination that considers two conditions.

- The first condition is distance, if the distance from a peak at *j-1* index to the peak at *j* index is bigger than the distance from peak at *j* to the peak at *j+1* index, score is increased one.
- Second condition by comparing local minima of the peak *j-1* and *j* index with the next local maxima of peak j and j+1 index. If a local minimum between j-1 index to j index is bigger than local minima of j to j+1, score is also increase one.
- If the peak passes two conditions, the score is automatically equal to 2, and this means that the peak is toe-off.

The proposed toe-off and heel-strike classification is a kind of 3 peaks sliding window that overlap 1 peak. Fig 3. is an algorithm for classifying toe-off and heel-strike.

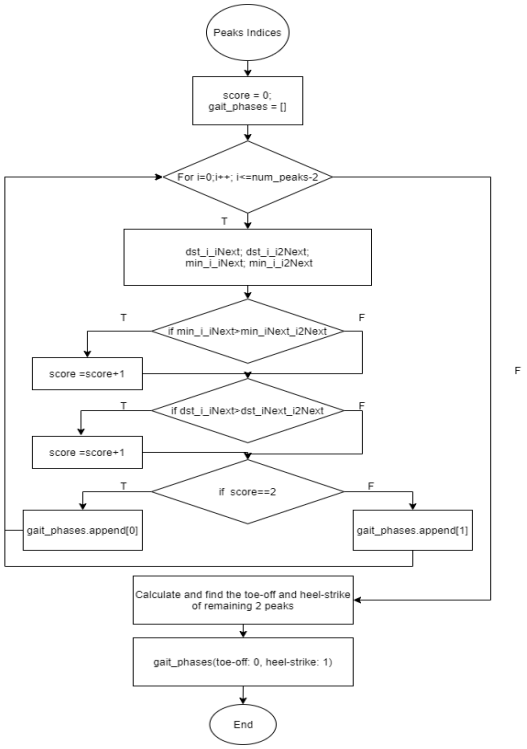



**Figure 5.** The proposed algorithm to classify gait phases.

### *D. Gait Parameters and Statistic Features Extraction*

The four gait parameters are calculated based on toe-off and heel-strike points. Stride time refers to the time between the heel-strike *k* index and the next heel-strike *k+1 index*. Stance time is the duration of foot contact on the ground. The swing time is the duration for which the foot leaves the ground foot in the air. Finally, the number of steps refers to the number of cycles the foot repeats the same action. The equations to calculate gait parameters are shown in Eq. (5)-(8).

$$Stride\_Time(k) = \frac{Heel\_Strike(k+1) - Heel\_Strike(k)}{Sampling\_Frequency} \quad (5)$$

$$Stance\_Time(k) = \frac{Toe\_Off(k) - Heel\_Strike(k)}{Sampling\_Frequency} \quad (6)$$

$$Swing\_Time(k) = \frac{Heel\_Strike(k+1) - Toe\_Off(k)}{Sampling\_Frequency} \quad (7)$$

$$Number\_Step = (Number\_Heel\_Strike - 1)/2 \quad (8)$$

We cannot extract gait parameters from the non-gait signal; Therefore, the statistic features are extracted from the non-gait signal. Those features include mean, standard deviation(std), and range shown in equation (9)-(11).

$$Mean = \frac{acc_x + acc_y + acc\_z}{3} \quad (9)$$

$$std = \sqrt{\sum (X - \bar{X})^2 / (n-1)} \quad (10)$$

$$Range = max(acc_x, acc_y, acc_y) - min(acc_x, acc_y, acc_z) \quad (11)$$

Where, acc_x, acc_y, and acc_z, are denote acceleration data along with the x, y, and z axis, respectively. std is the standard deviation of each individual data point, $\bar{X}$ is the mean of three acceleration points.

### *E. LSTM model*

The LSTM model comprises six layers: an input layer, four hidden layers, and an output layer. The input to the neural network is a sequence of gait parameters. For example, the input shape of (None, 251,3) has a fixed length sequence of 251 frames and three input features (gait parameters). This means there are 251 frames of sliding window and 125 overlapping every move to the next window. The output at the Dense layer is a binary number. It could be a value of 0 or 1. Design each model as follows: one input LSTM layer (units = 30), followed by three hidden LSTM layers (units = 60), one more hidden LSTM layer (unit = 30) and finally, one Dense output layer (units = 2, activation = softmax).

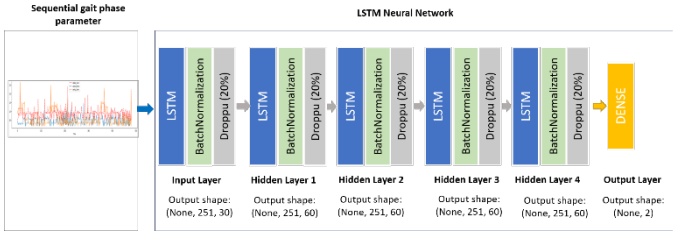



**Figure 6.** LSTM neural network architecture.

## IV. EXPERIMENT

All experiments were implemented on Windows 10 operating systems with Python version 3.9 and used the following libraries: Numpy version 1.23, Scipy version 1.10, Keras version 2.12, TensorFlow-gpu version 2.6.

### *A. Gait Phases Detection*

We used the *find_peaks* function of the Scipy Python library to detect local maxima and then proposed an algorithm to classify toe-off and heel-strike from those regional maxima indices. Peaks are detected by smoothing the y-axis multiplied by 10, with a height between 10.5 and 100, a distance of 30 and a width of 5.

### *B. Sequential Features*

Gait parameters and statistic features were extracted from gait, and non-gait signals, respectively. Then the extracted features from raw data were saved in a CSV file format where three columns(features) and raw is statistic features or gait parameter of each timestamp. The sliding window technique was applied for each file to conduct a sequence of data with a length of 250 samples. The sliding window moved a mong data file with a stride size of 125 overlapping samples. Finally, all sliding data from each file were combined. The total sliding

data is 22, 251, which were divided into 80%, 20% of training, and validation, respectively. Data from four patients (two high fall risk and low fall risk labels) were evaluated using the proposed fall risk assessment method, it contains 3,333 slides with a height is 250 and with of 3. The features extracted with sliding windows are shown in Fig. 7.


**Figure 7.** For example, in the features extracted from the signal of a high fall risk label, the blue signal represents the feature1(it could be stride-time or mean), the orange represents feature2 (it could be stance-time or standard deviation), the green color represents feature3(swing time or range values).

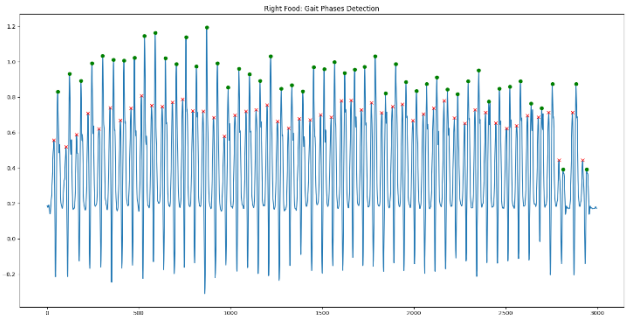
The developed LSTM model was trained with 100 epochs, Adam optimization with a learning rate of 0.0001, *binary_crossentropy* loss, and sigmoid activation function. In addition, *EarlyStopping* was used to avoid overfitting by monitoring validation loss with patience for 20 epochs.

## V. RESULTS

### A. Gait Phases Evaluation

The performance in detecting gait events like toe-offs and heel strikes was evaluated using 14 healthy people. The participants performed three tasks, walking 50 steps, 20 steps and fast walking 20 steps with free direction (turn left or turn right). The IMU sensors were placed on the ankle of both feet to capture the gait of walking. The IMU sensor contains 9 axes which are different from the IMU sensor used for collecting data from patients. Each of the participants walked into the multimedia building of Soonchunhyang University on the sixth floor. While walking, they need to count their steps manually to reach the target steps. There is one participant cannot walk fast.

The detection of gait phases on 50 steps walking from a single foot of a healthy person is presented in Fig. 8. And Fig. 9 and 10 present gait phases detection from patients with different risk levels.



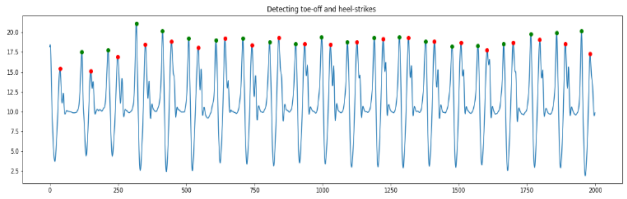




**Figure 9.** Example of detecting gait phases from segmented gait data into 20 seconds (low fall risk), red circles and green circles denote toe-offs and heel-strikes, respectively.


**Figure 10.** Detect gait phases from segmented gait data into 20 seconds (high fall risk), red circles and green circles denote toe-offs and heel strikes, respectively.

The mean percentage error (MPE) is used to evaluate the error of the gait phase detection. The $MPE_{sc}$ equation is mean percentage error of step counting.

$$MPE_{sc} = \frac{\sum_{i=1}^{np}(|NStep_{Act} - NStep_{Pred}|)/NStep_{Act}}{np} \times 100 \quad (12)$$

Where np denotes the number of participants, $NStep_{Act}$ denotes the number of actual step counts, and $NStep_{Pred}$ presents the number of step counts predicted by the proposed method.

The $MPE_{gp}$ equation is the mean percentage error of gait phase extraction. The equation is as follows:

$$MPE_{gp} = \frac{\sum_{i=1}^{np}(|Time_{Act} - Time_{Pred}|)/Time_{Act}}{np} \times 100 \quad (13)$$

Where $Time_{Act}$ represents the actual strid-time, stance-time, or swing-time; $Time_{Pred}$ represents the predicted duration of strid-time, stance-time, or swing-time using equations (5)-(7).

As the result of three experiments, walking 50 steps, 20 steps, and fast walking 20 steps the MPE is 0.43% error, 0.89% error, and 1.15% error, respectively. This error is very low. Therefore, the proposed gait phase detection is high performance with an accuracy of 99.57%, 99.11%, and 98.85 for walking 50 steps, walking 20 steps, and fast walking 20 steps, respectively. Table 1 presents the mean errors, standard deviation errors (Std-Error), MPEs (pean percentage errors), and performances for each experiment. Additionally, stride times, stance times, and swing times were extracted and compared with the actual times (time series recorded while walking). Table 2 evaluation performance and errors gait parameters compared with actual times.

**Table I.** Errors and performance of Step Counting for Each Experiment: Walking 50 steps, walking 20 steps, and Fast Walking 20 steps fast.

| | Mean Error | Std-Error | $MPE_{sc}$ | Accuracy |
|---|---|---|---|---|
| Walk 50 Steps | 0.0043 | 0.0098 | 0.43% | 99.57% |
| Walk 20 Steps | 0.0089 | 0.03 | 0.89% | 99.11% |
| Fast Walk 20 Steps | 0.012 | 0.021 | 1.15% | 98.85% |

**Table II.** Performance evaluation of proposed gait phase detection: Errors and Performances for Each Experiment.

| | Walk 50 steps | | Walk 20 steps | | Fast-Walk 20steps | |
|---|---|---|---|---|---|---|
| | MPE | Acc | MPE | Acc | MPE | Acc |
| Stride-Time | 5.08 | 94.92 | 5.40 | 94.60 | 4.65 | 95.35 |
| Stance-Time | 7.33 | 92.66 | 7.99 | 92.01 | 7.38 | 92.62 |
| Swing-Time | 7.92 | 92.08 | 8.23 | 91.77 | 7.62 | 92.38 |

B. Fall Risk Assessment

Right feet signal of four patients (2: high fall risk, and 2 of low fall risk level) were evaluated the performance of the proposed fall risk assessment method. The sliding window on the test signal with the same length as the train signal (250 samples). Therefore, the test signal contains samples: 3333, length: 250, and features:3. Among of test data, 1853 is the correct prediction as low fall risk and 428 is the correct prediction as high fall risk level. The proposed fall risk assessment method achieved an accuracy of 68%. The correct and incorrect prediction test data are shown in Fig. 11.

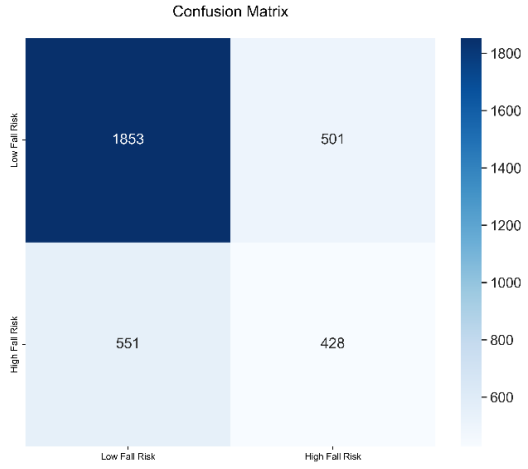



**Figure 11.** Confusion matrix of proposed method for classifying level of fall risk.

## VI. DISCUSSION AND CONCLUSIONS

This paper uses a deep learning model to classify low and high fall risk from long-term gait data. The gait and non-gait detection is completed with good detection. The proposed method to detect gait events such as toe-offs and heel strikes achieved better performance for healthy people. However, the performance of fall risk assessment needs to improve with current accuracy of 68%. In the future, we aim to improve the performance of the fall risk assessment method.

## ACKNOWLEDGMENT

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) program(IITP-2023-2020-0-01832) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation)

## REFERENCES

[1] N. Roth et al., "Real-World Stair Ambulation Characteristics Differ Between Prospective Fallers and Non-Fallers in Parkinson's Disease," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 9, pp. 4733-4742, Sept. 2022, doi: 10.1109/JBHI.2022.3186766.

[2] J. Wilczyński et al., "Assessment of risk factors for falls among patients with Parkinson's disease," in BioMed research international, vol.2021, pp.1-8, 2021.

[3] M. Ullrich et al., "Detection of Gait From Continuous Inertial Sensor Data Using Harmonic Frequencies," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 7, pp. 1869-1878, July 2020, doi: 10.1109/JBHI.2020.2975361.

[4] Y. C. Han, K. I. Wong and I. Murray, "Gait Phase Detection for Normal and Abnormal Gaits Using IMU," in IEEE Sensors Journal, vol. 19, no. 9, pp. 3439-3448, 1 May1, 2019, doi: 10.1109/JSEN.2019.2894143.

[5] L. Borzì et al., "Predicting Axial Impairment in Parkinson's Disease through a Single Inertial Sensor," Sensors, vol. 22(2), pp. 412, 2022.

[6] M. Sarshar, S. Polturi, & L. Schega., "Gait phase estimation by using LSTM in IMU-based gait analysis—Proof of concept," Sensors, vol. 21(17), pp. 5749, 2021.

[7] C. Tunca, G. Salur and C. Ersoy, "Deep Learning for Fall Risk Assessment With Inertial Sensors: Utilizing Domain Knowledge in Spatio-Temporal Gait Parameters," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 7, pp. 1994-2005, July 2020, doi: 10.1109/JBHI.2019.2958879.

[8] Ullrich, Martin, Nils Roth, Arne Küderle, Robert Richer, Till Gladow, Heiko Gaßner, Franz Marxreiter, Jochen Klucken, Bjoern M. Eskofier, and Felix Kluge. "Fall risk prediction in Parkinson's disease using real-world inertial sensor gait data." IEEE journal of biomedical and health informatics (2022).

[9] T. Van Hamme, G. Garofalo, E. Argones Rúa, D. Preuveneers, & W. Joosen., "A systematic comparison of age and gender prediction on IMU sensor-based gait traces," Sensors, vol. 19(13), pp. 2945, 2019

[10] J. Y. Kim, G. Park, S. A. Lee, & Y. Nam., "Analysis of machine learning-based assessment for elbow spasticity using inertial sensors," Sensors, vol. 20(6), pp. 1622, 2020.

[11] J. Y. Kim et al., "Gait analysis in patients with neurological disorders using ankle-worn accelerometers," The Journal of Supercomputing, vol. 77, pp. 8374-8390, 2020.

# Utilizing Textual Reviews to Predict the Age of Children by Text Mining Analysis with Machine Learning for the Language Fluency Tests

Nab Mat[1], Chomyong Kim[2], and Yunyoung Nam[3*]
[1] *Department of ICT Convergence, Soonchunhyang University, Asan, South Korea*
[2] *ICT Convergence Research Centre, Soonchunhyang University, Asan, South Korea*
[3] *Department of Computer Science and Engineering, Soonchunhyang University, Asan, South Korea*
*Contact: ynam@sch.ac.kr

***Abstract*— Accurate the age of the children from text analysis based on text mining with machine learning and deep learning is playing an important role not only in language fluency tests, but also in medication treatment, education, developmental milestones, and so on. Children experience a variety of problems with language fluency skills, such as language delay, speech sound disorder, learning disabilities, stuttering, and more troubles that are caused by genetics, hearing loss, developmental delay, brain injury, environmental factors, and lack of language stimulation. Therefore, the mentioned issues may have a negative impact on the ability to communicate effectively and may lead to irritation and trouble in society. In particular, since early detection and training are important for these symptoms, continuous observation is required, but the treatment period is often missed because there is a lack of related centres and there is no standard for parents to judge the child's language level. This paper proposes a linear regression-based prediction model for predicting the age of children based on their textual review using text mining and machine learning techniques. By analysing the pattern and features of the language in the reviews of children, our approach is able to accurately predict their age. Our approach can help educators and language professionals better understand the ability of individual children. The dataset will be collected from children aged between 6 to 12 years old from movies review. The research data and results of this study can confirm which adjectives are utilized by age group and will be utilized and contributed as basic data and related research areas on sentiment analysis by age group.**

## I. INTRODUCTION

The vast amount of data was generated through Internet use, the data was generated in the form of reviews, opinions, sentiments, and other forms that are required to extract the insights and intended information [1]. To meet the commercialized benefits, it is required to perform accurate models for the data analysis method. Furthermore, the elastic method that is suitable for large-scale and different datasets will contribute to the text mining analysis [2-3]. Text mining analysis is conducted to predict the age of the children based on their reviews from the experience of watching a movie. However, accurate model inference of the age of children predictions based on their reviewed comments are under significant challenges because there is insufficient demographic information.

Age-specific lexicons are relied on using age-specific lexicons can help researchers better understand how emotional expression changes as children develop [4]. In addition, to improve diagnosis and treatment, the analysis can potentially help diagnose emotional disorders and develop effective treatments. Furthermore, predicting the age of the children can be helpful for teachers and educators to understand the developmental stage of their students, and to find various strategies for studying accordingly. Especially, this approach represents a significant contribution to the field of language assessment and analysis, as it offers a new way to predict the age of children in language fluency tests.

Insight into child development gains contributions from analysing the ability of children with their writing skills and expression of their experience; researchers can gain insights into how the development of ability is related to other aspects of development, such as language acquisition, education, healthcare, child welfare, and cognitive development. Overall, the use of text mining analysis in this context can lead to a deeper understanding of children's emotional expression with writing skills of their age and provide useful insights for both research and clinical practice [5]. Moreover, we demonstrate the usefulness of our approach through experiments on language fluency tests, achieving quality performance in predicting the age of children in our dataset. Additionally, in this paper, we explore the potential applications of our approach in other domains where predicting age based on textual data is relevant, such as clinical research or education.

This paper presents a text mining analysis to predict the age of children based on their review after watching a movie to do the language fluency tests. A comparison

of age-specific lexicons that explore the use of text mining techniques to analyse the ability of the language fluency tests for children. The article presents a detailed analysis of the results and discusses the implications of the findings for language assessment and education. In addition to that, this paper highlights the potential of text mining and machine learning procedures to predict the age of children based on their textual reviews. By improving the high result of performance and fairness of language fluency tests, the out approach has the potential to benefit educators, language professionals and children around the world.

The rest of the paper is organized as follows. Section II presents the existing works which are included the relevant works in applied text mining analysis with applied machine

learning. Section III presents a details explanation of the dataset and how it is gathered, and the proposed method architecture is presented. In Section IV, the research results are presented and discussed. Section V includes the conclusions of the work and future research.

## II. Related Work

The existing literature on age prediction of children is based on their reviews by utilizing natural language processing (NLP) techniques with machine learning [5]. For instance, the utilization of convolutional neural networks (CNN) to predict the age of children based on their product reviews.

Furthermore, a deep neural network model combined text and image characteristics to predict the age of children based on their online activities and used a support vector regression (SVR) model to predict the age of students based on their course reviews [6].

The various BERT models on diversity NLP tasks were applied by Biseda el al., [9]. Text mining solutions were implemented to identify the context and full information of the combined data of the various articles [7-8]. In addition, the transformer model architectures for classification solutions on drug based on user reviews data was explored by Shiju et al., [10].

Zhou et al., [11] introduced the topic of distributed text mining systems for online web textual data analysis which presented a novelty of layered-based architecture for three layers for dedicated functions.

Garg [12] presented the recommendation system based on an applied machine learning algorithm utilizing text mining solutions, the work based on sentiment prediction which achieved 93% of the system accuracy metrics.

## III. Method

We have selected the review text and age of children from “Common Sense Media”, which is a website that provides reviews and ratings of movies, TV shows, books, games, podcasts, and others for parents and children. We scrapped children’s age and reviews and saved them to a CSV file, which contains a total of 32337 instances with two columns age and reviews texts. We designated children aged in between of 6 to 12 years old. Fig. 1 shows the proposed model used to predict the age of children based on their review. It carries seven stages, especially Data Collection, Data Cleaning, Data Pre-processing, Feature Extraction, Regression Model Prediction, and Model Evaluation.

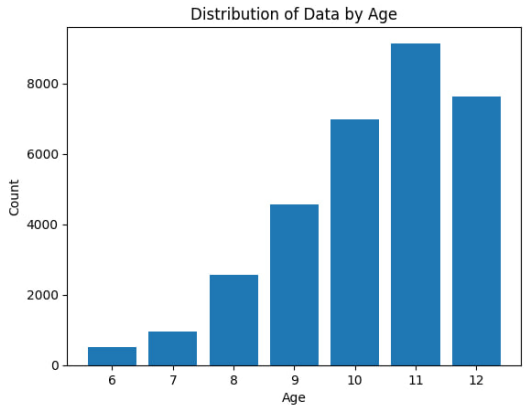



Fig 1. Comparison of age count in datasets

TABLE I
THE MIN/MAX NOUN, VERB, ADVERB, AND ADJECTIVE IN EACH AGES

| Age | Noun | Verb | Adjective | Adverb |
|---|---|---|---|---|
| | min/max | min/max | min/max | min/max |
| 6-year-old | 0/233 | 0/25 | 0/36 | 0/35 |
| 7-year-old | 0/201 | 0/16 | 0/27 | 0/28 |
| 8-year-old | 0/250 | 0/24 | 0/40 | 0/24 |
| 9-year-old | 0/391 | 0/49 | 0/35 | 0/39 |
| 10-year-old | 0/307 | 0/37 | 0/96 | 0/33 |
| 11-year-old | 0/447 | 0/41 | 0/50 | 0/58 |
| 12-year-old | 0/591 | 0/56 | 0/41 | 0/39 |

### A. Data Collection

We have used open-source Selenium, Beautifulsoup4 for scraping the children's reviews and ages. After that, we save it in a CSV file for our further processing or analysis. We selected the review of the children from various movies which is they explain about their experience on the movie they have watched.

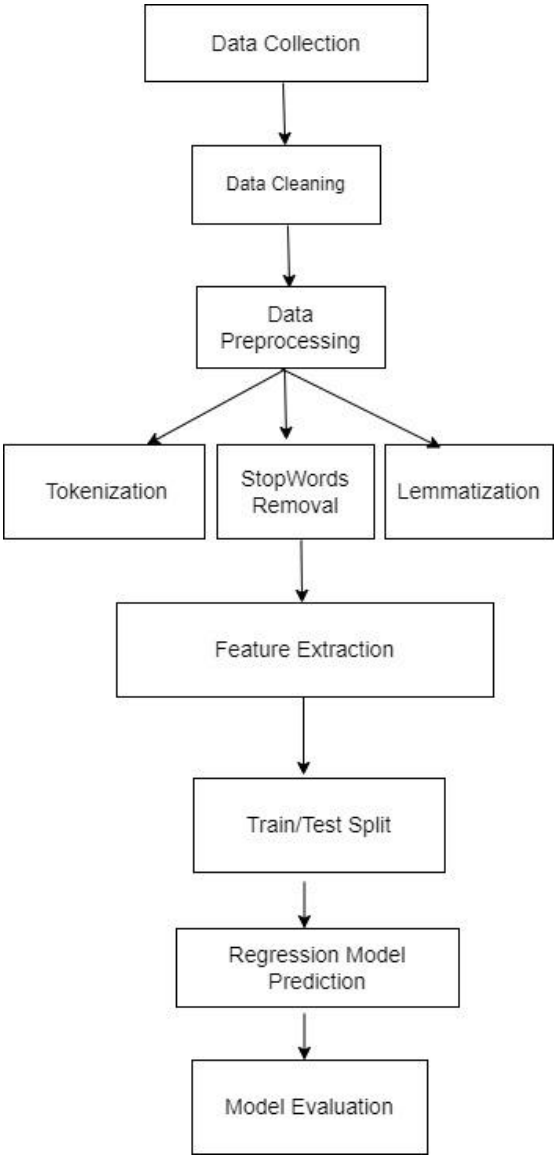



Fig 2. Flowchart of the proposed model

### B. Data Cleaning

Data cleaning is a significant step in the data analysis process and to enhance the quality of the model learning. It

involves identifying and correcting errors. In this stage, we performed the following steps, such as checking missing and dropping null values, checking, and removing emoji and special characters, removing the link, converting age to integer, removing punctuation, removing emoticons and non-printable words, and replacing all consecutive whitespace characters.

### C. Data Pre-processing

Data pre-processing plays the significant roles as main contributions in terms of improving model accuracy. The goal of the data pre-processing is improving the reliability of data and address any issues or consistencies insights from the data. At this stage, we convert text to lowercase, tokenization, Stopwords removal, and lemmatization. The computations of all parts which involve the cleaning and transforming of unstructured data to a well-structured format that can be easy for computers to understand, especially, and can be used for analysis.

Convert text to lowercase: this is the common way to convert text to lowercase. It can be consistent because all the words are in the same format. Moreover, when we convert text to lowercase, it reduces the number of unique words in a corpus.

Tokenization: it is the process of breaking down sentence text from a document into individual words. It is useful for text cleaning, it makes it easier to remove unwanted characters, such as punctuation, link, HTML tags, and so on.

Stopwords removal: Stopwords are common words in a language that occur frequently in the text but generally do not carry much meaning or value for text analysis. Removing stop words will contribute to lessening the amount of noise in the data. The first process is to custom the list of Stopwords that contained in the dataset. Next, we add Stopwords to the Stopwords list from nltk. Lastly, we remove the Stopwords.

Lemmatization: The lemmatization will contribute to lessening the inflectional forms of a word to a common base form, which can simplify text analysis and improve accuracy. Lemmatization is often used in conjunction with other pre-processing techniques like tokenization, Stopwords removal, and stemming.

### D. Feature Extraction

After data pre-processing, the important setup of the data is required for prediction and for computer understanding. Machine learning algorithms cannot work with text; therefore, we need to convert all text into numerical for processing and learning from the input data.

Bow: it stands for Bag of Words, which is commonly used in feature extraction. The BOW approach involves creating a vocabulary of all unique words in a corpus of text and then representative of any document in the amount as a vector of word frequencies.

Convert Vectorizer: Convert textual data into numerical vectors that can be used in machine learning algorithms.

Feature Vectors of Training: Training feature vectors are necessary to convert raw text data into a proper format that can be further processed by applying machine learning algorithms.

Feature Vectors of Testing: In machine learning, feature vectors of testing are used to represent the input data to test the performance of machine learning models that have already been trained on a set of feature vectors of training.

## IV. Evaluation and Result

### A. Model Evaluation

To compute the mean squared error, follow these steps:

1. Calculate the differences between the predicted and actual values for each data point.
2. Square each difference.
3. Compute the average of the squared differences.

The formula for MSE is:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

where:

$n$ is the number of data points.

$y_i$ is the actual value of the $i-th$ data point.

$\hat{y}_i$ is the predicted value of the $i-th$ data point.

To compute the mean absolute error, follow these steps:

1. Calculate the absolute difference between the predicted values and the actual values for each data point.
2. Compute the average of the absolute differences.

The formula for MAE is:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

where:

$n$ is the number of data points.

$y_i$ is the actual value of the $i-th$ data point.

$\hat{y}_i$ is the predicted value of the $i-th$ data point.

### B. Results

The machine learning regression model performed the result of prediction of children's age based on their review with different outcomes of each machine learning model, respectively. To evaluate the model of prediction, we compute the mean square error and mean absolute error which is common use to evaluate a regression model in machine learning. Each machine learning regression model performed the result, respectively; for Linear Regression with mean square error = 6.41 and mean absolute error = 1.25. Decision Tree Regression performed the result with mean square error = 1.28 and mean absolute error = 0.53. For the Support Vector Regression model, the result with mean square error =1.24 and mean absolute error = 0.77. The better regression model is the Random Forest Regression model, which is performed by mean Our performance outcome shows that our approach can accurately predict the age of children based on their reviews, with an error of approximately 10 months, and the last regression model which is Neural Network Regression performed result with mean square error =1.15 and mean absolute error = 0.58.

TABLE II
THE COMPARISON OF THE MODEL EVALUATIONS

| Model | Mean Square Error | Mean Absolute Error |
|---|---|---|
| Linear Regression | 6.41 | 1.25 |
| Decision Tree Regression | 1.22 | 0.51 |
| Random Forest Regression | 0.84 | 0.65 |

| | | |
|---|---|---|
| Support Vector Regression | 1.24 | 0.77 |
| Neural Network Regression | 1.15 | 0.58 |

## V. CONCLUSION

In this paper, we have presented a method for predicting the age of children based on their reviews. Our approach utilizes natural language processing techniques for data pre-processing and feature extraction from children's reviews. In addition, we predicted the age of the children by utilizing machine learning algorithms which are regression models to train a predictive model. Overall, our study is passionate about the potential of using natural language with machine learning to predict the age of the children based on their review. We hope that our approach can inspire further research in the area and lead to the development of more accurate and efficient predictive modes.

Moreover, to strengthen the research outcome, extensive experimentation and data preparation will be furthermore implemented. Additionally, the increased use of text mining and machine learning in language assessment continues to advance, we may see more widespread use of text mining and machine learning technique in language assessment, particularly for predicting the age based on textual data.

Furthermore, in the future, text mining and machine learning in language assessment is likely to be the important field in technology, as well as broader social and ethical considerations. As the field continue to evolve, it will be important for researchers and practitioners to stay abreast of development and work to ensure that their work is responsible, transparent, and equitable.

## ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2023-00218176).

## REFERENCES

[1] J. Breckling, Ed., The Analysis of Directional Time Series: Applications to Wind Speed and Direction, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[2] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," IEEE Electron Device Lett., vol. 20, pp. 569–571, Nov. 1999.

[3] M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, "High resolution fiber distributed measurements with coherent OFDR," in Proc. ECOC'00, 2000, paper 11.3.4, p. 109.

[4] R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, "High-speed digital-to-RF converter," U.S. Patent 5 668 842, Sept. 16, 1997.

[5] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.

[6] T. Matsumoto, W. Sunayama, Y. Hatanaka and K. Ogohara, "Data Analysis Support by Combining Data Mining and Text Mining," 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), Hamamatsu, Japan, 2017, pp. 313-318, doi: 10.1109/IIAI-AAI.2017.165.

[7] S. Binkheder, H. -Y. Wu, S. Quinney and L. Li, "Analyzing Patterns of Literature-Based Phenotyping Definitions for Text Mining Applications," 2018 IEEE International Conference on Healthcare Informatics (ICHI), New York, NY, USA, 2018, pp. 374-376, doi: 10.1109/ICHI.2018.00061.

[8] O. Kovalchuk, S. Banakh, M. Masonkova, K. Berezka, S. Mokhun and O. Fedchyshyn, "Text Mining for the Analysis of Legal Texts," 2022 12th International Conference on Advanced Computer Information Technologies (ACIT), Ruzomberok, Slovakia, 2022, pp. 502-505, doi: 10.1109/ACIT54803.2022.9913169.

[9] Biseda, Brent, and Katie Mo. "Enhancing Pharmacovigilance with Drug Reviews and Social Media." ArXiv.org, 18 Apr. 2020,arxiv.org/abs/2004.08731.

[10] Shiju, Akhil, and Zhe He. "Classifying Drug Ratings Using User Reviews with TRANSFORMER-BASED Language Models." MedRxiv, 15 Apr. 2021, doi:10.1101/2021.04.15.21255573.

[11] B. Zhou, Y. Jia, C. Liu and X. Zhang, "A Distributed Text Mining System for Online Web Textual Data Analysis," 2010 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, Huangshan, China, 2010, pp. 1-4, doi: 10.1109/CyberC.2010.11.

[12] S. Garg, "Drug Recommendation System based on Sentiment Analysis of Drug Reviews using Machine Learning," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021, pp. 175-181, doi: 10.1109/Confluence51648.2021.9377188.

# Binary Classification in Working Oil Data

Jaeeun Kim[1] and Mucheol Kim[1*]
[1]*Department of Computer Science and Engineering, Chung-Ang University, Seoul, Republic of Korea*
*Contact: kimm@cau.ac.kr

*Abstract*—**For construction equipment, it is important to understand the abrasion conditions and lubrication performance of internal mechanical parts. It determines the state of internal construction equipment parts whether normal or abnormal through condition of working oil. In this study, knowledge distillation, its variant models and anomaly detection methods are applied. As a result, the model using knowledge distillation shows relatively high performance and anomaly detection method shows low performance. The main conjecture of low performance is insufficient amount of data and class imbalance among labels. Although the overall performances are not very high, the performance could be improved to some extent by adjusting the model parameters.**

## I. Introduction

For construction equipment, it is important to understand the abrasion conditions and lubrication performance of internal mechanical parts. If the conditions can be identified through monitoring the internal working oil, it is possible to determine an appropriate replacement cycle and prevent accidents due to the abrasion.

This study used the data from the competition 'Construction Machine Oil Condition Classification AI Competition' [4] hosted by Hyundai Genuine and supervised by DACON. It aims to develop a model that performs binary classification condition whether normal or abnormal according to working oil condition (i.e. content of specific element). The number of available features at inference are limited compared to the features available for training.

## II. Background

Knowledge distillation (KD) is one of the model compression methods, which transfer the information from large model or ensemble of model, and trains a small model without a significant degradation of accuracy. In KD, large teacher model guides a small student model and the student model imitates the teacher's behavior [1]. To transfer the information of the large, complex model to a small model, a small model uses the class probabilities from complex model as a soft targets for the training. Since the soft targets have high entropy, they can provide much more information in training than hard targets. Therefore small student model can be trained on less data than the original teacher model with higher leraning rate [2].

## III. Methodology

### A. Teacher Assistant Knowledge Distillation

Teacher Assistant Knowledge Distillation (TAKD) model proposed by [3] uses teacher assistant (TA) model to alleviate the gap between teacher and student model. The intermediate size TA network is distilled from the teacher, and the TA behave like a teacher to student while it trains the student via distillation.

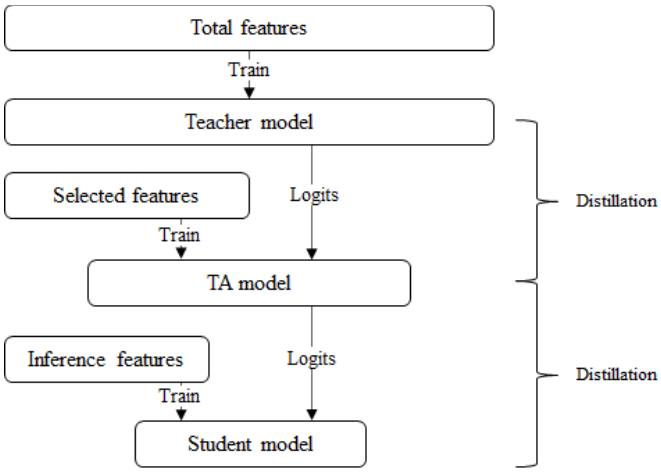



Fig. 1 Overview of Teacher Assistant Knowledge Distillation

In the Vanilla KD model, the distilled student model accuracy increases first but decreases with the increasing teacher size because the student does not have the sufficient capacity to imitate the complex teacher's behavior. The introduction of TA model can relieve the gap between teacher and student and therefore more effective fitting can be achieved to student model using the logit distribution of TA than the teacher's.

### B. Anomaly Detection

The fully-connected encoder is trained only using the normal labeled data in training step. The encoder takes total features of data and reduce the dimension. The anomaly score is computed as cosine similarity value of the results of encoder. The assumption is that the normal data is more aligned than the abnormal one and I considered higher similarity value as normal one. The abnormal data is considered as which has lower anomaly score than threshold value in inference step, which is the minimum value of anomaly score in training step which use only normal data as threshold value.

## IV. Experiments

### A. Dataset

The train data is concists of total 52 features and the label which deicde the normal or abnormal state. The test data is consists of limited 18 features used in diagnosis of oil conditions. The data imbalance between normal and

abnormal labels is large. There are some missing values according to the features.

TABLE I
STATISTICS OF DATASET

| | Features | Normal label | Abnormal label | Total |
|---|---|---|---|---|
| Train set | 52 | 12892 (91.5%) | 1203 (8.5%) | 14095 |
| Test set | 18 | - | - | 6041 |

### *B. Baseline models*

The baseline is selected as the vanilla KD which only uses the teacher and student model. Also two naïve methods using only inference features, which are fully-connected layers and support vector machine are selected as baseline model.

### *C. Teacher Assistant Knowledge Distillation*

The TA model is introduced between the teacher and the student model. In TAKD-based model, the selected features which used in the training of TA model (Figure 2) are same one as the second variant of chapter 4.3. The experiment is conducted using the TAKD-base model and variant of it. The variant model introduces one more TA model, named sub-TA between the TA and the student model. Sub-TA distilled from the TA model and also train the label from 34 numbers of randomly selected non-inferential features which are blue columns in Figure 3.

### *D. Anomaly Detection*

The fully-connected encoder with trainable parameters takes total features and reduces the dimension. Cosine similarity is calculated with randomly selected latent vector in minibatch. Since the imbalance of dataset, the mean value of cosine similarity will be close to 1. The training data only contains the normal set, therefore the minimum cosine similarity value is taken as the threshold value of anomaly score which has lower anomaly score than threshold can regarded as abnormal.

## V. RESULTS

The competition metric is macro f1 score which is the harmonic mean of the precision and recall. The Table 2 shows the competition score that the TAKD method achieved the highest f1 score. Vanilla KD and classification with fully connected layer achieved similar score respectively. TAKD with sub-TA model achieved less optimal performance than TAKD model. Anomaly detection and SVM method showed low performance than KD methods.

The conjecture of the result as follows: (1) the amount of data is not enough to apply more complicated methods (TAKD compared with vanilla KD or TAKD-based compared with Sub-TA) or anomaly detection. (2) the imbalance of label might cause the low performance of SVM.

TABLE II
EXPERIMENT RESULTS (F1 SCORE)

| Model | F1 Score |
|---|---|
| TAKD | **0.5651** |
| TAKD with Sub-TA | 0.5481 |
| Anomaly detection | 0.514 |
| Vanilla KD | 0.5523 |
| Fully connected NN | 0.5504 |
| SVM | 0.4844 |

## VI. CONCLUSIONS

The study about the binary classification in working oil data, which determines the state of internal construction equipment parts whether normal or abnormal through condition of working oil., is conducted in paper. Knowledge distillation, anomaly detection methods are used and each f1score is compared respectively. As a result, the model using KD shows relatively high performance than others, while the model using anomaly detection methods shows lower performance. The main conjecture of low performance is insufficient amount of data and class imbalance among labels. Although the overall performances are not very high, the performance could be improved to some extent by adjusting the model parameters.

## ACKNOWLEDGMENT

This research was supported in part by Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (P0012724, The Competency Development Program for Industry Specialist).

## REFERENCES

[1] Gou, J., Yu, B., Maybank, S. J., & Tao, D. (2021). Knowledge distillation: A survey. International Journal of Computer Vision, 129(6), 1789-1819.

[2] Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531, 2(7).

[3] Mirzadeh, S. I., Farajtabar, M., Li, A., Levine, N., Matsukawa, A., & Ghasemzadeh, H. (2020, April). Improved knowledge distillation via teacher assistant. In Proceedings of the AAAI conference on artificial intelligence (Vol. 34, No. 04, pp. 5191-5198).

[4] Construction Machine Oil Condition Classification AI Competition. DACON. (n.d.). Retrieved December 14, 2022, from https://dacon.io/competitions/official/236013

# Analysis of Wrist Photoplethysmography Data During Various Exercise Activities

Dayeong So[1], Subeen Leem[2], Insu Jeon[2], Yongsung Kim[3], and Jihoon Moon[1,2,4,*]
[1]*Department of ICT Convergence, Soonchunhyang University, Asan, South Korea*
[2]*Department of Medical Science, Soonchunhyang University, Asan, South Korea*
[3]*Department of Technology Education, Chungnam National University, Daejeon, South Korea*
[4]*Department of AI and Big Data, Soonchunhyang University, Asan, South Korea*
*Contact: jmoon22@sch.ac.kr, phone +82-41-530-4956

*Abstract*—**This study explored the Wrist Photoplethysmography (PPG) and Accelerometry data collected during walking, running, and cycling activities using simultaneous motion estimates collected from an accelerometer and gyroscope to reduce motion interference. A baseline chest electrocardiogram (ECG) was also included for comparison with the gold standard heart rate during exercise. Eight participants (three males and five females) aged between 22 and 32 years performed various types of exercise while wearing ECG, PPG, and inertial measurement devices. We used three datasets (high_resistance_bike, low_resistance_bike, and walk) of S1 participants for exploratory data analysis (EDA), where we visualized the data using line plots. Our analysis revealed no significant difference in the ECG pattern between bike and walk activities, while the PPG data showed that the walk had a larger amplitude than the bike. We plan to leverage distributed systems and cloud computing technologies to analyze and visualize the vast amount of data effectively in future studies.**

## I. INTRODUCTION

The field of medicine is changing rapidly, and collaboration and data-driven approaches have become increasingly important [1]. Advances in information technology have made remote healthcare more accessible and efficient, leading to the generation of vast amounts of data [2]. Innovative methods are necessary to analyze and interpret these massive datasets to optimize health outcomes [3], [4]. In bioinformatics, extracting valuable knowledge from these datasets in the era of big data is increasingly important. Data analytics is an effective method for analyzing complex biomedical data, including electrocardiogram (ECG) signals [5]. ECG data can provide valuable insights into heart health, and analyzing these signals effectively is essential for improving patient care and optimizing health outcomes. By using data analytics and other techniques, such as machine learning, it is possible to gain insights into heart health and develop personalized healthcare approaches. These techniques are crucial in bioinformatics as they can contribute to advancements in medical research and the development of novel treatments [6]. By leveraging the insights gained from these methods, healthcare professionals can make informed decisions and optimize patient care [6], [7]. Therefore, using data-driven approaches in healthcare is crucial for improving the quality of healthcare services and outcomes [8].

This study aims to analyze ECG data collected during various exercise activities to gain insights into heart health. ECG data can provide valuable information about the heart's electrical activity, and abnormalities in this activity can indicate underlying heart conditions [9]. Effective analysis of ECG data can lead to the early detection and monitoring of such conditions, resulting in improved patient care and health outcomes [10]. In addition to ECG data, we also analyze data from the Wrist Photoplethysmography (PPG) During Exercise dataset. PPG is a non-invasive optical method that measures changes in blood volume, providing information about the cardiovascular system [11]. The PPG data can complement the ECG data by providing additional insights into the cardiovascular system and helping in the analysis and interpretation of the ECG data [11], [12]. We aim to better understand heart health during exercise activities through data visualization and exploratory ECG and PPG data analysis. The findings of this study can provide valuable insights for healthcare professionals and researchers in developing personalized healthcare approaches and improving health outcomes.

The insights gained from this analysis can provide valuable information to stakeholders, including healthcare professionals and policymakers. By understanding the latest solutions impacting public health, stakeholders can make informed decisions and develop effective strategies to improve health outcomes. Moreover, these insights can help academics understand the current research needs in bioinformatics and biomedical research. Section 2 defines the relevant terms and concepts used in the analysis, including ECG and PPG, to ensure a comprehensive study understanding. This section also provides necessary background information on the methods and techniques used in this study. In Section 3, we visualize the data from the Wrist PPG During Exercise dataset and discuss our findings. The data visualization and exploratory analysis results provide insights into heart health during exercise activities, which can inform the development of personalized healthcare approaches and improve patient care. Finally, Section 4 concludes the study and provides recommendations for future research. Ultimately, this study can significantly impact improving healthcare services and individual outcomes.

## II. Dataset Description

Section 2 provides a detailed description of the dataset used in the study. The dataset includes wrist PPG data recorded during walking, running, and cycling activities. Simultaneous motion estimates were collected using an accelerometer and gyroscope to reduce motion interference, offering multiple filtering options. A baseline chest ECG was also compared with the gold standard heart rate during exercise.

Participants in the experiment had an ECG device attached to their chest and a PPG and inertial measurement device attached to their left wrist. They used an indoor treadmill and exercise bike simultaneously. ECG recordings were taken using self-adhesive silver-silver chloride (Ag/AgCl) electrodes and an Actiwave recorder. The R peaks of the ECG traces were manually identified and included in the database for comparison with gold standard heart rate data. A Shimmer 3 GSR+ unit was used for PPG and motion data, integrating a gyroscope, low-noise accelerometer, wide-range accelerometer, and magnetometer into a single package. The PPG sensor was attached to the main Shimmer unit to provide a stable connection. The combined unit was worn on the left wrist instead of a standard watch.

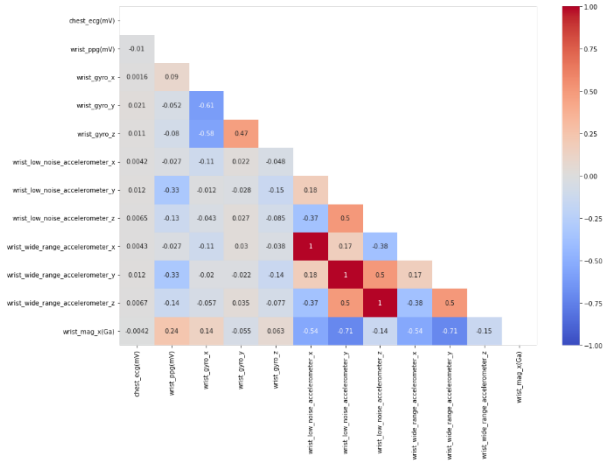
Participants performed different types of exercise, including walking, light jogging/running, cycling with low resistance, or cycling with high resistance, to introduce representative motion artifacts into the cardiac signals. Participants could adjust the speed and resistance and stop the exercise anytime. Most participants spent between 4 and 6 minutes on each activity. The dataset includes records from eight participants (three males and five females), aged between 22 and 32 years (average 26.5 years).

The walking and running records in the database contain raw PPG and motion signals that appear as an appropriate activity after segmentation. No filtering is applied beyond what is built into the Shimmer hardware. In the case of cycling recordings, a significant amount of high-frequency noise was present in the PPG traces. Before converting to WFDB format, cycling PPG traces were low-pass filtered using a second-order IIR Butterworth digital filter with a 15 Hz cutoff and zero group delay with Matlab. All ECG records had a 50 Hz notch filter applied as part of the Actiwave control software to remove main interference.

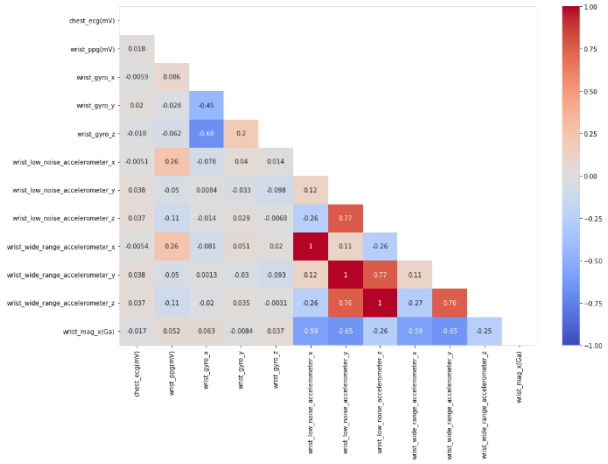
## III. Exploratory Data Analysis

In Section 3, we provide a clear understanding of the Wrist PPG During Exercise dataset and its relevant concepts and data used in our analysis. The dataset was obtained from Soonchunhyang University's big data-based living lab education platform (BLEP).

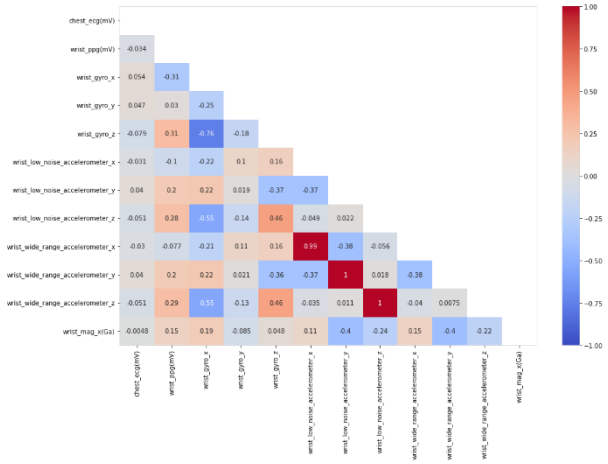
We used three datasets from S1 participants for exploratory data analysis (EDA), which included various types of sensor data, such as ECG, PPG, gyroscope, low noise accelerometer, wide range accelerometer, magnetometer, and measurement time. These variables were recorded numerically and arranged into columns for each dataset. For example, the *chest_ecg(mV)* variable represents the ECG data, while the *wrist_ppg(mV)* variable represents the PPG data. The *wrist_gyro* variable corresponds to the gyroscope data, while the *wrist_low_noise_accelerometer* and *wrist_wide_range_accelerometer* variables correspond to the low-noise accelerometer and wide-range accelerometer data, respectively. The *wrist_mag(Ga)* variable represents the magnetometer data, while the *sample_times_for_all_signals_apart_from_ecg(s)* variable presents the measurement time for all signals apart from the ECG data in seconds. Fig. 1 shows the correlation matrix for each dataset, including the high_resistance_bike, low_resistance_bike, and



(a) *high_resistance_bike*

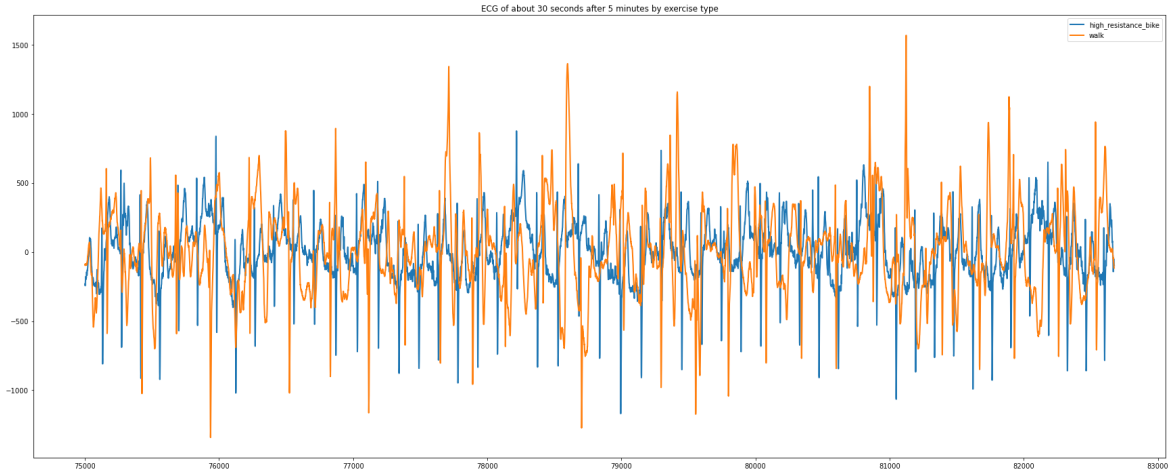


(b) *low_resistance_bike*

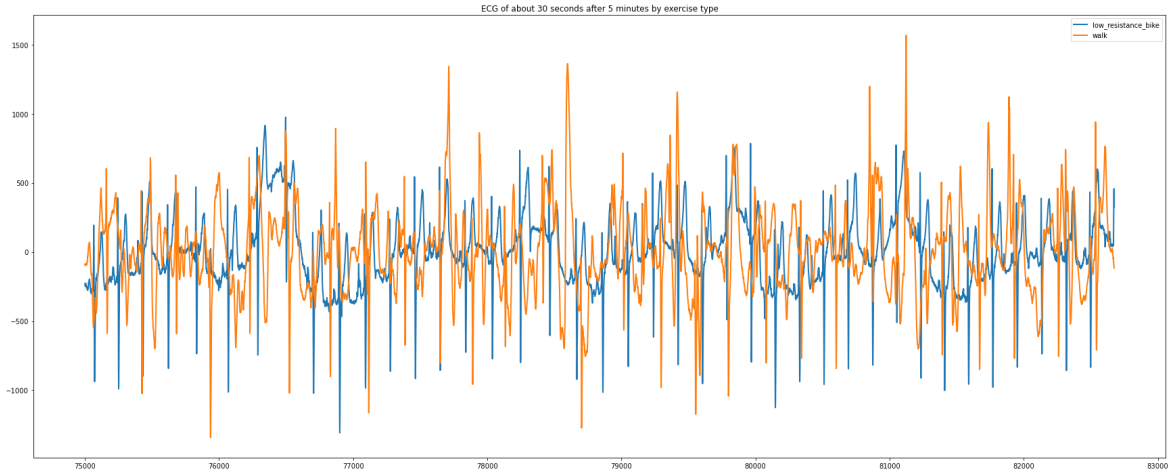


(c) *walk*

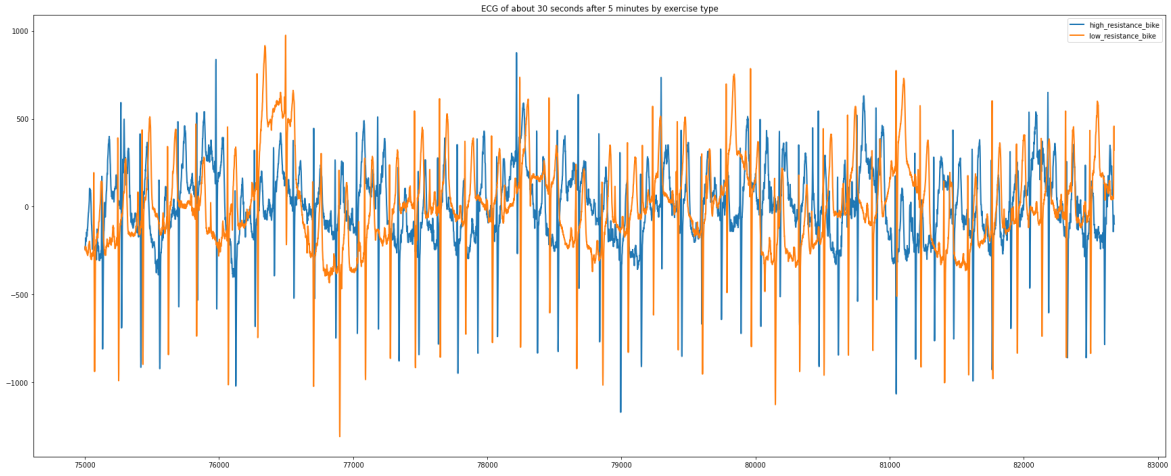
Fig. 1 Correlation matrix for each data set




(a) *high_resistance_bike* (blue) vs. *walk* (orange)




(b) *low_resistance_bike* (blue) vs. *walk* (orange)




(c) *high_resistance_bike* (blue) vs. *low_resistance_bike* (orange)

Fig. 2 ECG line plot between two variables

walk. The matrix represents the correlation coefficients between each pair of variables in the dataset. The darker the color, the stronger the correlation between the variables. The matrix provides a visual representation of the relationships between the variables in each dataset, allowing us to identify any patterns or trends in the data.

We used line plot visualizations to analyze the data in the *high_resistance_bike*, *low_resistance_bike*, and *walk* datasets of S1 participants. Our analysis of the ECG line plot in Fig. 2 showed almost no difference in the bike and walk patterns, contrary to our initial expectation. In the PPG line plot in Fig. 3, we confirmed that the walk had a larger amplitude than the

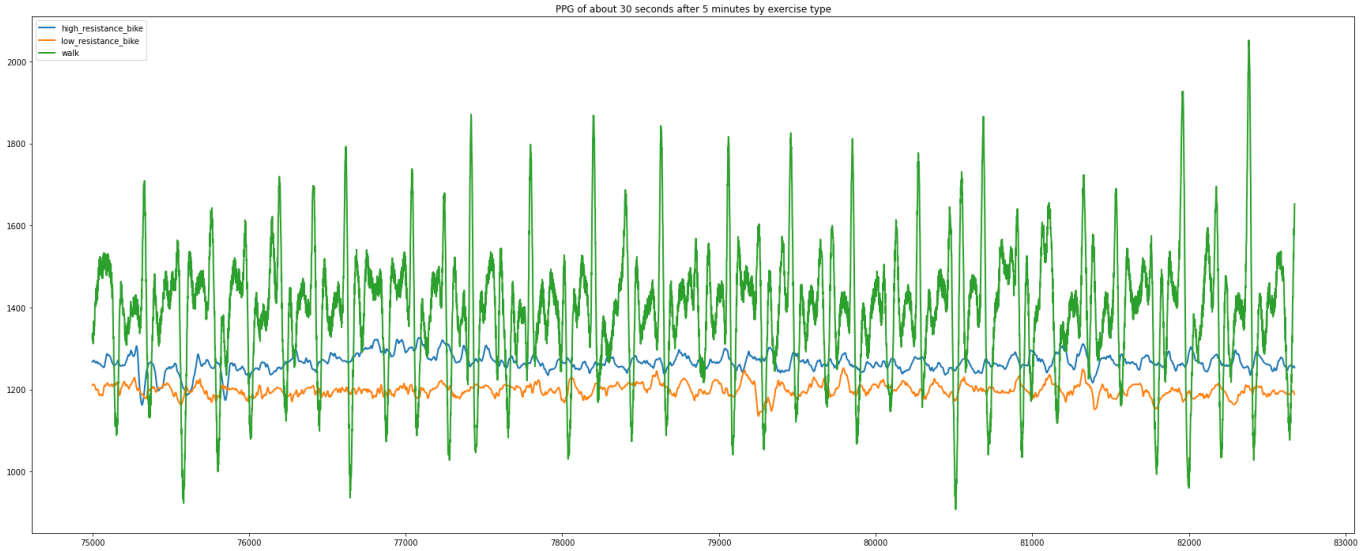



Fig. 3 PPG line plot between high_resistance_bike (blue), low_resistance_bike (orange), and walk (green)

bike from the viewpoint of both activities, and there was no difference in the pattern of the high and low data of the bike. However, analyzing and visualizing such massive data was complex and challenging. Therefore, in the future, we plan to use distributed systems and cloud computing technologies to construct EDA and machine learning models effectively.

## IV. Conclusions

This study analyzed the Wrist PPG During Exercise dataset, which includes ECG and PPG data collected during walking, running, and cycling activities at both high and low levels. The dataset involved eight participants, and we performed EDA on three datasets to gain valuable insights into the data. Our analysis of the ECG line plot showed almost no pattern difference between the bike and walk, contrary to our initial expectations. In the PPG line plot, we confirmed that the walk had a larger amplitude than the bike, and there was no difference in the pattern of the high and low data of the bike. We used line plot visualizations to aid our analysis but faced complexity due to the massive amount of data.

Our study uses PPG and ECG data to provide valuable insights into heart health during exercise. These findings have significant implications for developing wearable devices that monitor cardiovascular health during exercise. Through analyzing data from multiple participants engaged in various exercise activities, our study offers a comprehensive understanding of the Wrist PPG During Exercise dataset, highlighting the importance of exploratory data analysis in analyzing complex biomedical data.

In the future, we plan to utilize distributed systems and cloud computing technologies to construct EDA and machine learning models effectively. Our study provides necessary definitions and background information on relevant concepts and data used in the analysis, ensuring a comprehensive understanding of the study's methods and techniques. The results of this study contribute to a better understanding of the Wrist PPG During Exercise dataset and provide a basis for further research in this field.

## Acknowledgment

This work was supported by the Soonchunhyang University Research Fund.

## References

[1] G. Samuel and A. M. Lucassen, "The environmental sustainability of data-driven health research: A scoping review," *Digital Health*, vol. 8, 2022, Art. no. 20552076221111297.

[2] A. Banerjee, C. Chakraborty, A. Kumar, and D. Biswas, "Emerging trends in IoT and big data analytics for biomedical and health care technologies," in *Handbook of Data Science Approaches for Biomedical Engineering*, 2020, pp. 121-152.

[3] W. Fan and A. Bifet, "Mining big data: current status, and forecast to the future," *ACM SIGKDD Explor. Newslett.*, vol. 14, no. 2, pp. 1-5, 2013.

[4] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, "Big data in healthcare: management, analysis and future prospects," *J. Big Data*, vol. 6, no. 1, pp. 1-25, 2019.

[5] S. Min, B. Lee, and S. Yoon, "Deep learning in bioinformatics," *Brief. Bioinform.*, vol. 18, no. 5, pp. 851-869, 2017.

[6] D. V. Dimitrov, "Medical internet of things and big data in healthcare," *Healthc. Inform. Res.*, vol. 22, no. 3, pp. 156-163, 2016.

[7] Y. Wang, L. Kung, and T. A. Byrd, "Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations," *Technol. Forecast. Soc. Change*, vol. 126, pp. 3-13, 2018.

[8] D. Gotz and D. Borland, "Data-driven healthcare: challenges and opportunities for interactive visualization," *IEEE Comput. Graph. Appl.*, vol. 36, no. 3, pp. 90-96, 2016.

[9] M. J. M. Cluitmans, R. L. M. Peeters, R. L. Westra, and P. G. A. Volders, "Noninvasive reconstruction of cardiac electrical activity: update on current methods, applications and challenges," *Neth. Heart J.*, vol. 23, pp. 301-311, 2015.

[10] D. Swancutt, R. Hobbs, D. Fitzmaurice, J. Mant, E. Murray, S. Jowett, et al., "A randomised controlled trial and cost effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in the over 65s (SAFE) [ISRCTN19633732]," *BMC Cardiovasc. Disord.*, vol. 4, pp. 1-11, 2004.

[11] S. S. Mousavi, M. Firouzmand, M. Charmi, M. Hemmati, M. Moghadam, and Y. Ghorbani, "Blood pressure estimation from appropriate and inappropriate PPG signals using A whole-based method," *Biomed. Signal Process. Control*, vol. 47, pp. 196-206, 2019.

[12] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Front. Public Health*, vol. 258, 2017.

# Stability Analysis of Network Controlled Helicopter Systems with a Periodic Sampling Times

Seok Young Lee[1,*]
[1]*Electronic Engineering, Soonchunhyang University, Asan, Republics of Korea*
*Contact: suk122@sch.ac.kr

*Abstract*— **This paper is concerned with the stability analysis problem of a network-controlled helicopter system with a periodic sampling time. Due to inherent limit of hardware resources, networked control system has sampling times when controlling remote systems. By utilizing an input-delay approach, sampled-data system can be reformulated into a linear time-varying discrete-time system. Since this paper considers a periodic sampling case, the reformulated discrete-time system can be regarded as a linear time invariant discrete-time system. When this system has all eigenvalues inside a unit circle, the system is asymptotically stable. A novel numerical example which considers a practical helicopter system is discussed.**

## I. Introduction

This paper is concerned with the stability analysis problem of a network-controlled helicopter system via an input-delay approach. Between two sequential sampling times, a sampled-data system is controlled by previously sampled states and thus can be modelled as a time-delay system with a saw-tooth delay. Inevitably, information at the sampling instants have played essential roles in controlling the systems and reducing the conservatism of the stability criteria. This paper utilizes the results of [2] for stability analysis of a network-controlled helicopter system. A novel numerical example for a practical helicopter system demonstrates the effectiveness of [2] in terms of allowable sampling intervals.

## II. Main Results

Consider the following sampled-data system of [1].

$$\dot{x}(t) = Ax(t) + A_d x(t_k), \ \forall t \in [t_k, t_{k+1}), \quad (1)$$

where $x(t) \in R^n$ is the system state, $A, Ad \in R^{n\times n}$ are system matrices, and $t_k$ is the sampling instant such that $\cup k \in N[t_k, t_{k+1}) = [0, +\infty)$.

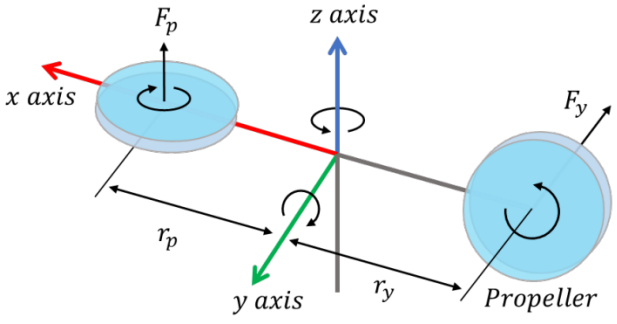



Fig. 1 A free-body diagram of the 2-DOF helicopter system

The sampling interval is defined as

$$t_{k+1} - t_k = h_k \quad (2)$$

Based on Theorem 1 and Theorem 2 of the paper [2], this paper newly discusses the Quanser AERO 2-DOF helicopter system [3] with a periodic sampling time. This helicopter system can be modelled as a free body diagram in Figure 1. This system consists of two identical rotors that produce the thrust forces $F_p(t)$ and $F_y(t)$ acting at two points with distances $r_p$ and $r_y$ from the z-axis along the x-axis, respectively. Thus, a propeller generates a torque around the y-axis leading to a pitch motion, while the other handles a yaw motion. This system can be described as follows:

$$\tau_p(t) = J_p\ddot{\theta}(t) + D_p\dot{\theta}(t) + K_{sp}\theta(t), \quad (3)$$

$$\tau_y(t) = J_y\ddot{\phi}(t) + D_y\dot{\phi}(t), \quad (4)$$

where $J_p = J_y = 0.0215$ are moments of inertia about pitch axis and yaw axis, respectively. $D_p = 0.0071$ and $D_y = 0.0220$ are viscous friction constants about pitch axis and yaw axis, respectively. $K_{sp} = 0.0374$ is the stiffness about the pitch axis. The torques $\tau_p(t)$ and $\tau_y(t)$ which respectively acts on the pitch and the yaw axes are assumed to be proportional to the input DC voltages Vp(t) and Vy(t) of the rotors such that:

$$\tau_p(t) = K_{pp}V_p(t) + K_{py}V_y(t), \quad (5)$$

$$\tau_y(t) = K_{yp}V_p(t) + K_{yy}V_y(t), \quad (6)$$

where $K_{pp} = 0.0011$ and $K_{py} = 0.0021$ are thrust torque gain acting on pitch axis from pitch propeller and yaw propeller, respectively. Also, $K_{yp} = -0.0027$ and $K_{yy} = 0.0022$ are thrust-torque gains acting on yaw axis from pitch propeller and yaw propeller, respectively. By utilizing the equations (4)-(6) and the state variable $x(t) = [\theta(t)\ \varphi(t)\ \dot{\theta}(t)\ \dot{\varphi}(t)]T$ , this helicopter system can be represented as a sampled-data system (1) with the following matrices

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -K_{sp}/J_p & 0 & -D_p/J_p & 0 \\ 0 & 0 & 0 & -D_y/J_y \end{bmatrix}, \quad (7)$$

$$A_d = BK,\ B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ K_{pp}/J_p & K_{py}/J_p \\ K_{yp}/J_y & K_{yy}/J_y \end{bmatrix}, \quad (8)$$

$$K = \begin{bmatrix} 0.0432 & 0.0530 \\ 1.1617 & -0.6085 \\ -0.1687 & -0.2070 \\ -0.1789 & 0.0937 \end{bmatrix}^T . \quad (9)$$

Here, a matrix $K$ is a gain of a sampled-state feedback controller, and $B$ is a system matrix. In this example, i newly derive an analytic upper bound $h = 12.0942$ of a periodic sampling interval. With a periodic sampling time, sampled-data system can be regarded as a linear discrete-time system. Integrating the differential equation (1) yields

$$x(t) = \Gamma(t - t_k), t \in [t_k, t_{k+1}],$$
$$\Gamma(\tau) = e^{A\tau} + \int_0^\tau e^{A(t-\tau)} drA\_d\,, \tau \geq 0.$$

Under the periodic sampling $h = h_k$, the system dynamics in (1) becomes

$$x(t_{k+1}) = \Gamma(h)x(t_k) \quad (10)$$

This system is asymptotically stable if and only if $\Gamma(h)$ has all eigenvalues inside the unit circle. Fig. 2 demonstrates that the system with the periodic sampling interval $h_k \leq h = 12.0942$ is asymptotically stable. In Figure 2, all state responses with an initial condition $x(0) = [10\ 45\ 0\ 0]T$ converge.

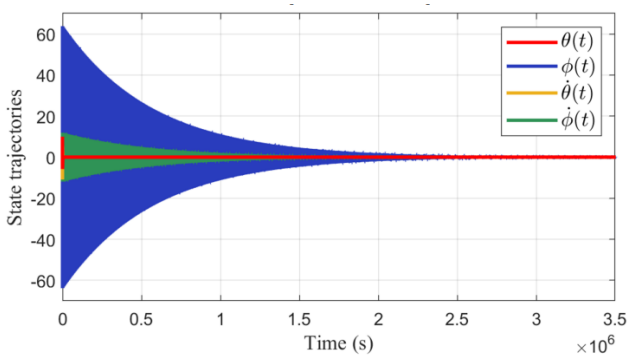



Fig. 2 The state $x(t) = [\theta(t)\ \varphi(t)\ \dot{\theta}(t)\ \dot{\varphi}(t)]T$ trajectories of the helicopter system with a periodic sampling interval $h_k = 12.0942$.

## III. Conclusions

This paper has discussed stability analysis of the sampled data controlled helicopter system with a periodic sampling time. In the future works, the derived results also can be applied to problems concerned with sampled-data controller synthesis and synchronization of various systems including fuzzy systems, switched systems, and delayed chaotic Lur'e systems.

## References

[1] N. K. Kwon and S. Y. Lee, "Novel equalities for stability analysis of asynchronous sampled-data systems," IEEE Access, vol. 8, pp. 177 195–177 205, 2020.

[2] S. Y. Lee, "Improved stability criteria for sampled data systems via a novel looped-functional," in The 13th Asian Control Conference (ASCC), May 2022.

[3] J.M. Park, "An improved stability criterion for networked control systems with a constant transmission delay," Journal of the Franklin Institute, vol. 359, no. 9, pp. 4346–4365, 2022.

# Efficacy of acute resistance exercise on cognitive function in older adults

Jin-Hyuck Park[1*]

[1]Department of Occupational Therapy, College of Medical Science, Soonchunhyang University, Asan

*Abstract*—**Benefits of acute resistance exercise with moderate intensity on cognitive function in older adults are still up for debate. This study looked into the effects of acute resistance exercise in healthy older adults. 24 healthy older individuals were randomly assigned to the experimental group (EG) which engaged in moderately intense resistance activity or the waitlist control group (CG) for a total of 15 sessions. To compare both groups, the Trail-Making Test Part B (TMT-B) was used to assess executive function, and prefrontal cortex activity during executive function testing was analyzed. Following the intervention, the EG achieved a significant improvement in the TMT-B ($p < 0.001$), compared to the CG. Additionally, the EG demonstrated a lower HbO2 concentration in the prefrontal cortex during the TMT-B ($p < 0.05$) than the CG. These findings provide new insight into the clinical efficacy of acute resistance exercise with moderate intensity on executive function and prefrontal cortex activation in healthy older adults.**

## I. Introduction

With normal aging, cognitive decline frequently manifests as impairments in executive function based on the prefrontal cortex (PFC) [1]. It has been demonstrated that executive dysfunction brought on by aging results in attentional dysregulation, which impairs general cognitive function in older persons [2]. As a result, numerous cognitive therapies have been explored to slow down cognitive decline.

Physical exercise has drawn a lot of attention among these cognitive therapies due to its positive effects on cognitive processes, particularly executive function [3]. Chronic physical activity has been shown to alter the brain's structural makeup, including increasing the amount of gray and white matter and improving cognitive function [4,5]. These encouraging results can be linked to the fact that physical activity raises arousal levels in direct proportion to its intensity [3,6]. The stimulation of neurophysiological channels into the PFC during exercise promotes optimal cognitive performance [6]. The majority of earlier research to date has been on the long-term effects of moderate to high-intensity exercise on cognitive function and has found distinctly beneficial effects [6,7]. The fact that the intensity of exercise is thought to be inversely connected with adherence to a long-term period of exercise, however, supports the concept that older persons find it challenging to maintain moderate to high-intensity exercise for a long time [8]. Recent research on the effects of acute exercise on cognitive function has begun to draw increasing interest from a practical standpoint [6]. Acute exercise, in particular, may be more advantageous for older persons who have sedentary lifestyles [9]. Acute exercise does, in fact, have a significant impact on cognitive performance in older persons, according to a prior study.3 Additionally, a number of neuroimaging studies with a focus on the PFC have shown neural substrates of acute exercise, confirming its beneficial benefits [2,3,7].

Functional near-infrared spectroscopy (fNIRS), an optical technique that monitors cerebral hemodynamics in a non-invasive manner by measuring changes in the attenuation of near-infrared light passing through tissues, has been introduced in a number of studies [2,11]. Dorsolateral PFC (DLPFC) activity and neural substrates for acute exercise-induced increased executive function in older adults have been successfully proven in prior fNIRS investigations [2, 3, 6]. However, acute resistance exercise was only undertaken in the majority of fNIRS research in order to examine its fast effects, and the effects of a series of acute resistance exercise performed for training purposes were not examined [2,3,6].

This study used fNIRS to investigate if acute resistance exercise of moderate intensity can also have a positive impact on executive function and the neural bases of these functions in the DLPFC. The study's main hypothesis was that moderately intense acute resistance exercise significantly improves executive function and stimulates DLPFC activity in relation to executive function.

## II. Methods

### A. Design

A single-blinded, randomized controlled trial was used in this study, which was carried out at a senior center in Seoul, South Korea. The wait-list control group (CG), which participated in acute exercise training after this trial as did the experimental group (EG), was randomly assigned to receive the acute resistance exercise training program with moderate intensity (Fig. 1). A person who was not participating in the study used computer-generated random numbers to carry out the randomization. Pre- and post-intervention outcome measures were carried out by a blinded assessor with six years of clinical experience.

### B. Participants

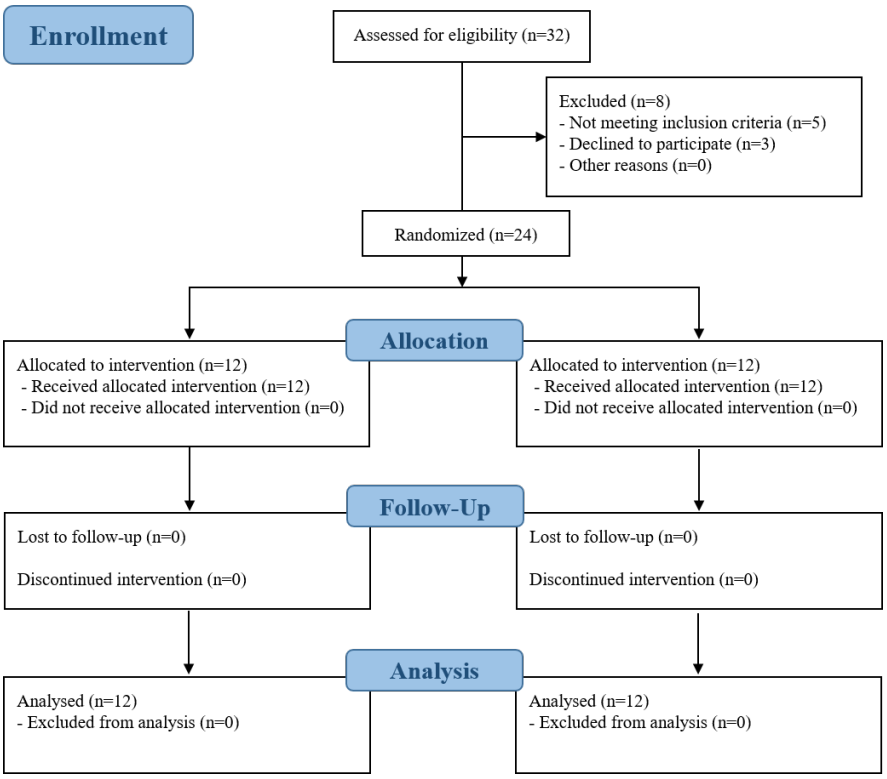



Fig. 1. Flow diagram of the study process

In this study, 32 healthy older persons living in local community in Seoul, South Korea, were recruited via senior centers. Subjective cognitive deterioration was not reported by any subject. If a participant completed the following requirements, they were all considered eligible and enrolled in the study: A disease requiring medical attention, the ability to exercise continuously for 30 minutes at a moderate intensity, and evidence of intact global cognitive function, as measured by the Korean version of the Mini-Mental Status Examination (MMSE-K) score of 24, among other criteria. The following were the exclusion criteria: (a) the presence of visual or auditory impairments; (b) the use of any pharmacological or non-pharmacological cognitive therapy during the previous three months; and (c) a history of heart disease, severe hypertension, or arthritis. All subjects provided their written consent.

### C. *Intervention*

An occupational therapist with 9 years of professional experience led the resistance exercise. To decide the exercise intensity, 1 repetition maximum was measured. Moderate-intensive resistance exercise was performed at 30% 1RM (12 repetitions) for 15 training sessions. The exercise routine consisted of three sets of five exercises (leg extension, leg curl, squat, seated row, and arm curl) for each exercise and a rest ratio of 1:2 [13, 14].

### D. *Outcome measures*

The Trail-Making Test Part B (TMT-B) was used in this study to evaluate the executive function specific to the PFC function (TMT) [17,18]. 25 circles are dispersed across a sheet of paper in the TMT-B. There are numerals and letters in these circles. The instructions required the participants to join the rings in ascending and alternate number-letter order. The TMT-B was given a maximum of 300 seconds to finish, and the overall completion time was recorded [18].

Using the OctaMon fNIRS instrument (Artinis, Netherlands), the PFC activity was captured. The near-infrared light used by this gadget has two wavelengths, 760 and 850 nm. The OctaMon can monitor hemodynamic responses such as oxygenated hemoglobin (HbO2) and deoxygenated hemoglobin (HHb) in the PFC and has a total of eight channels with light sources and detectors that are 3.5 cm apart. The light source and detector pairs in this investigation were positioned above the left (Fp1) and right (Fp2) frontal cortical areas, targeting the left and right dorsolateral PFC (DLPFC), in accordance with the modified international electroencephalography (EEG) 10-20 system [19].

The subjects were asked to clean their foreheads with alcohol swabs before putting on the fNIRS device and to keep their fringe up to prevent noise recording. The sampling frequency used for all data was 10 Hz [19]. The PFC activity was measured in this study while carrying out the TMT-B at the pre- and post-test phase. Instead of using HHb to measure PFC activity, which is sensitive to assessing cognitive function, HbO2 concentration was used [20].

### E. *Statistical analyses*

SPSS 22.0 version was used to analyze all the data. Utilizing SPSS, version 22.0 (SPSS Inc., USA), all data were examined.

The general features of the subjects were presented using descriptive statistics. The Kolmogorov-Smirov statistic was used to determine whether all continuous variables had a normal distribution. The Chi-square and t-test were used to find significant baseline differences in both groups' general characteristics and outcome variables. Following the fifteen training sessions, a dependent t-test was used to compare both groups in the outcome measures. The cutoff for statistical significance was $p < 0.05$.

## III. Results

### A. General characteristics of the participants

Following the inclusion and exclusion criteria, a final selection of 32 suitable individuals from 61 healthy senior citizens was made (Figure 1). Table I shows that there were no statistically significant differences in the respondents' demographics ($p$'s > 0.05). Between the two groups, there were no considerable differences in the characteristics (Table I).

TABLE I

DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF BOTH GROUPS

| Characteristics | | Experimental group (n = 12) | Control group (n = 12) | $\chi^2$ / t |
|---|---|---|---|---|
| Demographic characteristics | | | | |
| Sex | Male | 7 (53.8%) | 6 (50.0%) | .168 |
| | Female | 5 (41.7%) | 6 (50.0%) | |
| Age (years) | | 69.08 ± 4.71 | 71.58 ± 3.17 | 1.523 |
| Stroke type | Infarction | 6 (50.0%) | 7 (58.3%) | .168 |
| | Hemorrhage | 6 (50.0%) | 5 (41.7%) | |
| Onset period (years) | | 9.50 ± 2.61 | 9.08 ± 2.10 | .430 |

### B. Executive function

There was a significant difference in the TMT-B ($p < 0.001$) (Table II). Specifically, the EG subjects showed a shorter reaction time for the TMT-B than the CG subjects.

### C. Prefrontal cortex activity

There was a significant difference in DLPFC activity, showing lower HbO2 in the prefrontal cortex while performing the TMT-B ($p < 0.05$) (Table II).

## IV. Discussion

Acute exercise's effects on executive function and DLPFC activation in healthy older individuals were examined. According to the findings, after 15 training sessions, acute moderate exercise improved executive function and reduced DLPFC activity during executive functioning, which is in line with earlier research [2-4,7,21].

Even cognitive deficits associated with lowered working memory, inhibition, and long-term memory are part of normal aging [1]. As a result, numerous earlier research has attempted to examine the efficacy of therapies to stop or enhance cognitive decline in healthy older persons [22]. Resistance exercise has been shown to be one of these therapies that helps healthy older individuals' cognitive performance [22]. Recent research specifically shows that resistance exercise has the best effects on executive function [23]. These findings are consistent with the findings of this investigation and suggest that resistance exercise would be therapeutically advantageous in older persons without cognitive impairment. According to the current study, those in the EG who underwent resistance exercise of moderate intensity displayed significantly improved executive function compared to those in the CG. An increasing amount of research has shown that regular physical activity significantly improves cognitive function in healthy older persons [2-4,7,22,23]. The TMT-B employed in this study is recognized to have a strong correlation with executive function [24]. Working memory, interference inhibition, mental flexibility, and

| | Experimental group (n = 12) | Control group (n = 12) | Between-group differences | P |
|---|---|---|---|---|
| TMT-B (sec) | | | | |
| Pre-intervention | 129.14 ± 7.01 | 127.36 ± 8.58 | | |
| Post-intervention | 124.11 ± 9.37 | 127.21 ± 8.19 | 4.89 | <.001 |
| Within-group changes | 2.66 ± 1.96 (1.41; 3.91)** | 0.14 ± 1.62 (-0.31; 1.97) | | |
| DLPFC activity (mm/mol) | | | | |
| Pre-intervention | 1.021 ± 0.125 | 1.137 ± 0.225 | | |
| Post-intervention | 0.931 ± 0.090 | 1.083 ± 0.182 | 0.036 | <0.05 |
| Within-group changes | 0.089 ± 0.060 (3.03; 5.56) *** | 0.053 ± 0.056 (0.34; 2.15) * | | |

problem-solving abilities are often components of executive function using a top-down mental process that underlies goal-directed tasks [25]. In particular, the improvement in TMT-B performance after the resistance exercise could be linked to executive function because the TMT-B can assess a skill related to working memory, which mostly depends on executive function [2,26].

Resistance exercise has been shown to be positively related to executive function, as it is closely associated with increased volumes in the PFC and its function represented as executive function [4]. The aforementioned results were further supported by earlier neuroimaging research that showed resistance exercise could increase brain activity in the PFC by increasing neuronal recruitment in the region [4]. The cardiorespiratory fitness hypothesis states that increased maximal oxygen consumption following resistance exercise would increase regional cerebral blood flow, resulting in a better oxygen supply to the PFC than other brain areas since the PFC is the area of the brain where age-related cerebral blood flow decline is most pronounced [27]. However, in terms of its quick effects, these results have been shown very away after resistance exercise. On the other hand, this study supported the benefits of moderately intense resistance exercise [2,3,6]. The EG subject's executive function improved after 15 sessions of resistance exercise, and there was a decrease in HbO2 in the DLPFC during executive functioning. Training-induced neural efficiency may be interpreted as decreased brain activity accompanied by high performance on cognitive tests, meaning that the EG participants were able to reach higher levels of executive function with less mental effort than the CG ones [28]. This study's findings are supported by earlier research that found that following cognitive-related training, brain activity decreased and cognitive function improved [28,29]. This study demonstrates that relative long-term resistance exercise may improve neuronal efficiency in the DLPFC in healthy older individuals by decreasing its activity during executive function. According to earlier research, resistance exercise boosted levels of brain-derived neurotrophic factors, which in turn improved neuronal plasticity and led to the growth of new neurons [4]. As a result, the outcomes of this study could be supported by the brain being more effective and adaptive, which translated into greater cognitive function following resistance exercise. On the other hand, the compensation theory-based neuroimaging model suggests that healthy older persons may have PFC hyperactivity [27]. According to this concept, in order to offset the cognitive deficits associated with aging, older persons must maintain higher levels of neuronal activation [27]. The CG subject's performance in this investigation is consistent with this model.

## V. Conclusion

The results of this study showed that moderately intense acute resistance exercise could be useful in enhancing PFC executive function and neural efficiency. The recent findings have significant clinical significance because the neural substrate underlying the benefits of acute resistance exercise of moderate intensity employed as cognitive training has very rarely been explored. The use of an fNIRS instrument to research the effects of resistance exercise as cognitive training has also been made clearer by this work. Future research must assess the effects of acute resistance exercise at various intensities utilizing an fNIRS device with more channels in order to offer older person better options based on their physical condition.

## References


[1] D.C. Park, G. Lautenschlager, T.Hedden, et al. Models of visuospatial and verbal memory across the adult life span. Psychological Aging 17 (2002), 299-320.

[2] K. Hyodo, I. Dan, K. Suwabe, et al. Acute moderate exercise enhances compensatory brain activation in older adults. Neurobiological Aging 33 (2012), 2621-2631.

[3] K. Byun, K. Hyodo, K. Suwabe, et al. Positive effect of acute mild exercise on executive function via arousal-related prefrontal activations: An fNIRS study. Neuroimage 98 (2014), 336-345.

[4] S.J. Colcombe, A.F. Kramer, K.I. Erickson, et al. Cardiovascular fitness, cortical plasticity, and aging. Proceedings of the National Academy of Sciences 101 (2004), 3316-3321.

[5] K.I. Erickson, M.W. Voss, R.S. Prakash, et al. Exercise training increases size of hippocampus and improves memory. Proceedings of the National Academy of Sciences 108 (2011), 3017-3022.

[6] H. Yanagisawa, I. Dan, D. Tsuzuki, et al. Acute moderate exercise elicits increased dorsolateral prefrontal activation and improves cognitive performance with Stroop test. Neuroimage 50 (2010), 1702-1710.

[7] H. Budde, A. Brunelli, S. Machado, et al. Intermittent maximal exercise improves attentional performance only in physically active students. Archives Medical Research 43 (2012), 125-131.

[8] K.L. Cox, V. Burke, T.J. Gorely, et al. Controlled comparison of retention and adherence in home- vs center-initiated exercise interventions in women ages 40–65 years: The S.W.E.A.T. Study (Sedentary Women Exercise Adherence Trial). Preventive Medicine 36 (2003), 17-29.

[9] S. Rovio, I. Kåreholt, E.L. Helkala, et al. Leisure-time physical activity at midlife and the risk of dementia and Alzheimer's disease. Lancet Neurology 4 (2005), 705-711.

[10] C.H. Hillman, K.I. Erickson, A.F. Kramer. Be smart, exercise your heart: Exercise effects on brain and cognition. Nature Review Neuroscience 9 (2008), 58-65.

[11] H. Obrig, A. Villringer. Beyond the visible-imaging the human brain with light. Journal of Cerebral Blood Flow Metabolism 23 (2003), 1-18.

[12] C.S. Carter, A.M. Macdonald, M. Botvinick, et al. Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. Proceedings of the National Academy of Sciences 97 (2000), 1944-1948.

[13] H. Chang, K. Kim, Y.-J. Jung, et al. Effects of a moderate-intensity aerobic exercise programme on the cognitive function and quality of life of community-dwelling elderly people with mild cognitive impairment: A randomised controlled trial. International Journal of Nursing Study 93 (2019), 97-105.

[14] M.E. Nelson, W.K. Rejeski, S.N. Blair, et al. Physical activity and public health in older adults: Recommendation from the American college of sports medicine and the American Heart Association. Medicine Science Sports Exercise 116 (2007), 1094-1105.

[15] Y. Mori, T. Tobina, K. Shirasaya, et al. Long-term effects of home-based bench-stepping exercise training on healthcare expenditure for elderly Japanese. Journal of Epidemiology 21 (2011), 363-369.

[16] G. Borg. Borg's perceived exertion and pain scales. Champaign, IL: Human Kinetics, 1998.

[17] M.S. Olson, H.N. Williford, D.L. Blessing, et al. The cardiovascular and metabolic effects of bench stepping exercise in females. Medicine Science Sports Exercise 23 (1991), 1311-1317.

[18] J. Stroop. Studies of interference in serial verbal reactions. Journal of Experimental Psychology 18 (1935), 643-662.

[19] T.N. Tombaugh. Trail Making Test A and B: normative data stratified by age and education. Archives of Clinical Neuropsychology 19 (2004), 203-214.

[20] I. Maidan, H. Bernad-Elazari, N. Giladi, et al. A. When is higher level cognitive control needed for locomotor tasks among patients with Parkinson's disease? Brain Topography 30 (2017), 531-538.

[21] G. Strangman, D.A. Boas, J.P. Sutton. Non-invasive neuroimaging using near-infrared light. Biology Psychiatry 52 (2002), 679-693.

[22] A.T. Harveson, J.C. Hannon, T.A. Brusseau, et al. Acute effects of 30 minutes resistance and aerobic exercise on cognition in a high school sample. Res Quarterly Exercise Sport 87 (2016), 214-220.

[23] J.M. Northey, N. Cherbuin, K.L. Pumpa, et al. Exercise interventions for cognitive function in adults older than 50: A systematic review with meta-analysis. British Journal of Sports Medicine 52 (2018), 154-160.

[24] R.S. Falck, J.C. Davis, J.R. Best, et al. Impact of exercise training on physical and cognitive function among older adults: A systematic review and meta-analysis. Neurobiology of Aging 79 (2019), 119-130.

[25] M.P. Milham, M.T. Banich, A. Webb, et al. The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. Brain Research Cognition Brain Research 12 (2001), 467-473.

[26] S. Colcombe, A.F. Kramer. Fitness effects on the cognitive function of older adults: A meta-analytic study. Psychological Science 14 (2003), 125-130.

[27] K. Arbuthnott, J. Frank. Trail Making Test, Part B as a measure of executive control: Validation using a set-switching paradigm. Journal of Clinical Experimental Neuropsychology 22 (2000), 518-528.

[28] R. Cabeza, N.D. Anderson, J.K. Locantore et al. Aging gracefully: compensatory brain activity in high-performing older adults. Neuroimage 17 (2002), 1394-1402.

Y.Y. Liao, H.Y. Tseng, Y.J. Lin, et al. Using virtual reality-based training to improve cognitive function, instrumental activities of daily living and neural efficiency in older adults with mild cognitive impairment. European Journal of Physical Rehabilitation Medicine 56 (2020), 47-57.

# Electronic Medical Record-Based Machine Learning Model for Predicting Cardiovascular Disease Risk within 3 Years: Development and Validation

Myeongcheol Lee,[1] Hyeonjin Kim,[1] Hayeon Lee,[1] Sang Youl Rhee[1,2*]

[1]*Center for Digital Health, Medical Science Research Institute, Kyung Hee University Medical Center, Kyung Hee University College of Medicine, Seoul, South Korea,* [2]*Department of Endocrinology and Metabolism, Kyung Hee University School of Medicine, Seoul, South Korea*

**Correspondence: Sang Youl Rhee (rheesy@khu.ac.kr)*

*Abstract*— **We aim to develop and validate machine learning models that utilize structured and unstructured electronic medical record data to predict cardiovascular disease risk. The use of diverse data fields and external validation improves upon previous models and may lead to more accurate risk prediction tools. We used data from three medical centers in South Korea to develop machine learning algorithms for predicting cardiovascular disease risk in patients with diabetes mellitus. Four popular machine learning algorithms were employed, and the developed models were validated externally in two different hospitals. The study found that combining multiple categories of patient data, including demographics, medical history, and laboratory values. XGB exhibited the highest performance in the development cohort (AUROC of 0.85 (95% CI 0.83-0.86)), while RF had the highest performance in the external validation cohort (AUROC of 0.73). We developed and validated machine learning models for predicting cardiovascular disease risk using electronic medical record data, demonstrating reasonable performance and potential for real-world clinical application.**

## I. Introduction

Machine learning (ML) has become an increasingly utilized branch of artificial intelligence (AI) in the field of cardiovascular medicine.[1] Essentially, it involves the process by which computers process data and make decisions or classifications regarding a particular task, either with or without human supervision.[2] The fundamental framework of ML is based on models that receive input data, such as images or text, and use mathematical optimization and statistical analysis to predict outcomes, such as favorable, unfavorable, or neutral. Various ML algorithms have been applied in daily activities. For instance, support vector machines (SVMs) are commonly used to recognize non-linear patterns in facial recognition, handwriting interpretation, and detection of fraudulent credit card transactions. Boosting algorithms, which are used for prediction and classification, have been utilized to identify and process spam email. Another algorithm, known as random forest (RF), facilitates decision-making by averaging multiple nodes. Additionally, convolutional neural network (CNN) processing, which combines several layers, is used for image classification and segmentation.[3]-[5]

Recent years have seen an increase in the availability of large volumes of electronic medical record (EMR) data, leading to growing interest in data-driven approaches for constructing efficient risk prediction tools.[6], [7] Machine learning (ML) algorithms have emerged as a promising alternative to traditional models for risk prediction, owing to their ability to handle complex EMR data and overcome the limitations of previous models.[8] Studies on models based on EMR data for predicting adverse events have suggested that incorporating ML may improve the accuracy of risk prediction.[9] However, validated and robust models remain limited, as previous models have relied mainly on prespecified variables based on traditional risk factors, consisting primarily of structural data, and have lacked proper external validation. Therefore, the objective of this study was to develop ML models that utilize diverse fields of both structured and unstructured EMR data to predict the risk of cardiovascular disease, and to validate the model using a different cohort.

## II. Methods

### A. Database

The data utilized in this study were obtained from Kyung Hee Medical Center, a tertiary medical care facility that serves the South Korean population. This centre has a bed capacity of approximately 1,000 and receives around 940,000 outpatient clinic visits and 210,000 admissions annually across 31 departments. The registry used in this study included a total of 68,009 patients who met the inclusion criteria of being admitted as inpatients or visiting as outpatients for diabetes mellitus or seeking care in the emergency department for established or suspected diabetes mellitus between 2008 and 2022. Of the patients included in the registry, 34,433 (50.6%) were male and 33,576 (49.4%) were female. The study protocol has been endorsed by Kyung Hee University.

### B. External validation set

To evaluate the performance of our model in a diverse population, we conducted external validation using data from two different tertiary hospitals: Kyung Hee University Hospital at Gangdong, with around 1,000 beds, and Gachon University Gil Medical Center, with around 1,500 beds. The patient characteristics, medical practices, and electronic medical record systems vary between these hospitals, providing an opportunity to assess the generalizability of our model.

*C. Data processing*

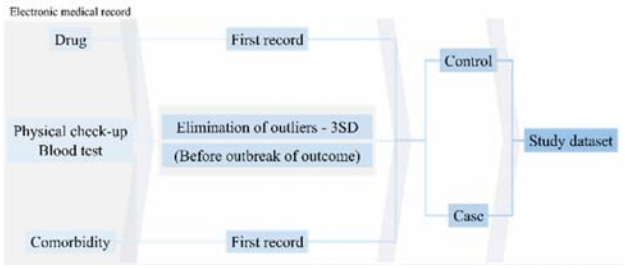
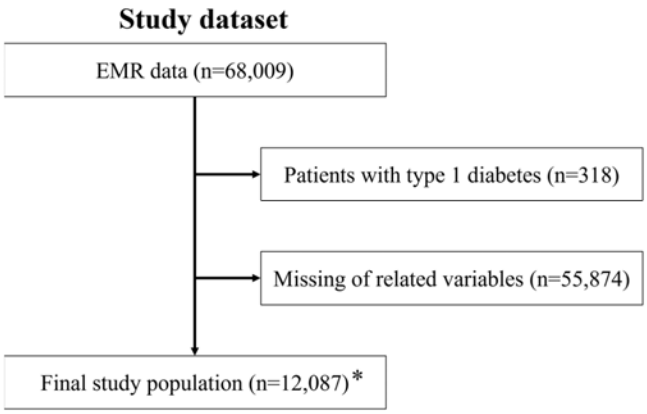



Fig. 1 The process of pre-processing EMR data



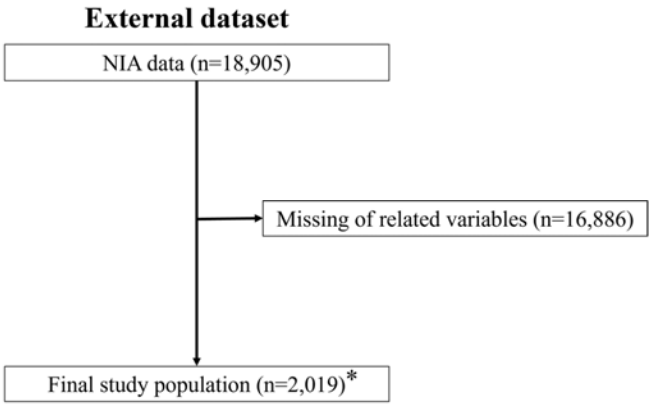




Fig. 1 Study workflow

Initially, we obtained anonymized records of 68,009 patients who had a history of diabetes mellitus and had visited one of three medical centers, namely Kyung Hee Medical Center, Kyung Hee University Hospital at Gangdong, or Gachon University Gil Medical Center. To eliminate errors and duplications, we established clinically plausible criteria. We then performed text mining to structure the significant variables associated with cardiovascular diseases, as most of the results of principal cardiovascular disease-related medical examinations are in free-text format. The text mining approach employed a three-step process. Firstly, we partitioned the readings into three frames, namely text, tabular, and others, and specified the extraction rules for each frame. Secondly, we constructed new tables by extracting the keywords and features from the original data. The values of the keywords were based on the rules defined in the preceding step. To enable interoperability for convergent multi-centre research, we standardized the data by using various codes corresponding to the common data model. Classic pre-processing technologies were used to obtain most of the structured data, including data cleansing, data integration, data transformation, data reduction, and privacy protection. Finally, we extracted structured data elements comprising demographics, administrative information, medical history and comorbidities, diagnoses, vital signs, laboratory values, and medications. We identified three categories of data using the developed algorithms, namely basic static data (demographics and administrative data), physical check-up and blood test data (physical data, laboratory values, and vital signs), and disease-specific data (medications, medical history, and comorbidities) (refer to Fig. 1 and Fig. 2).

Data collection and preparation were authorized by the institutional review board of Kyung Hee University Hospital, and the requirement for informed consent was waived. Patient deidentification was carried out in compliance with the Health Insurance Portability and Accountability Act. This report adheres to the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis reporting guideline.[10] All preprocessing were performed using the SAS software (version 9.4; SAS Institute, Cary, NC, USA).

*D. Machine learning algorithms*

We employed four widely used classes of machine learning (ML) algorithms in our study, namely extreme gradient boosting (XGB), random forest (RF), ada-boost (AB), and light gradient boosting machine (LGBM). XGB is an ensemble ML algorithm that utilizes the gradient boosting framework to construct a strong predictive model by combining multiple weak models.[11] It also implements parallel computing and tree pruning techniques to optimize execution speed and model performance.[12] RF is another ensemble ML algorithm that creates a forest of models using decision trees and combines their predictions to enhance accuracy and stability.[13] AB is a boosting algorithm that

TABLE I

HYPERPARAMETERS AND THEIR VALUES IN EACH MODEL

| | Booster | Learning rate | Max depth | N estimators | Subsample | Max features | Estimator | Colsample bytree | Num leaves |
|---|---|---|---|---|---|---|---|---|---|
| XGB | GBtree | 0.1 | 20 | 250 | 0.39 | N/A | N/A | N/A | N/A |
| RF | N/A | N/A | 10 | 150 | N/A | sqrt | N/A | N/A | N/A |
| AB | N/A | 0.5 | N/A | 50 | N/A | N/A | Decision tree classifier (max depth=2) | N/A | N/A |
| LGBM | N/A | 0.2 | 8 | 200 | 0.6 | N/A | N/A | 1 | 40 |

iteratively trains weak models and adjusts their weights based on their performance to form a robust predictive model that is resistant to overfitting.[14] LGBM, on the other hand, is a gradient boosting framework that employs a unique gradient-based one-side sampling technique to achieve faster training speed and lower memory usage, making it suitable for large-scale datasets and high-dimensional feature spaces.[15]

To optimize the performance of each model, we determined the hyperparameters using an empirical search and 10-fold cross-validation on the study population. Table 1 summarizes the hyperparameters and their respective values for each model. The optimal tuning parameter values were selected based on the testing accuracy values that were calculated for each fold and averaged. Furthermore, we validated the developed prediction models externally using a cohort from a different hospital. We also evaluated the performance of each data category and assessed the cumulative performance with combinations of multiple information categories by gradually adding each category one by one to identify the best performance. We utilized the Python programming language with the Keras library and Tensorflow backend to develop the risk algorithms in the training cohort and apply them to the validation cohort.

The study population's descriptive characteristics are presented in terms of percentages (%) for categorical variables and means with standard deviations (SD) for continuous variables. The discrimination performance of each model was assessed by calculating the area under the receiver operating characteristic curve (AUROC). Statistical significance was set at a 2-sided P value of less than 0.05.

## III. RESULTS

### *A. Baseline characteristics*

The characteristics of the study population were analyzed and presented in Table 2 for both the development and internal validation groups. The average age of the patients was 62.5 years with a standard deviation of 10.2 years. Among the 12,087 patients, 51.0% were male. The external validation cohort comprised 2,019 patients, of which 53.8% were male. The detailed characteristics of the patients in the external validation cohort were also presented in Table 2.

### *B. Performance*

The performance of the machine learning algorithms in predicting cardiovascular disease is presented in Table 3. Among the algorithms, XGB achieved the highest predictive performance with an AUROC of 0.85 (95% CI 0.83-0.86; P<.001). Both RF and LGBM demonstrated similar performance, with AUROCs of 0.83 (95% CI 0.82-0.85, P<.001) and 0.84 (95% CI 0.82-0.85, P<.001), respectively. However, the AB algorithm exhibited a lower AUROC of 0.80 (95% CI 0.78-0.81, P<.001).

The external validation of the ML models for predicting cardiovascular disease was conducted using datasets from Kyung Hee University Hospital at Gangdong and Gachon University Gil Medical Center. As shown in Table 4, RF demonstrated the highest predictive performance with an AUROC of 0.73, followed by LGBM with an AUROC of 0.71. XGB and AB showed slightly lower AUROCs of 0.70 and 0.69, respectively.

## IV. CONCLUSIONS

Our study aimed to develop and validate machine learning models for predicting cardiovascular disease risk by utilizing diverse fields of electronic medical record (EMR) data. The results showed reasonable performance of the ML algorithms, indicating that this approach can be generalized for various healthcare prediction models. We suggest that this study provides a substantial method for developing risk prediction models using real-world clinical data sets. Further studies are necessary to establish the clinical effectiveness of this approach and its applicability in real-time at the point of care.

## V. ACKNOWLEDGMENTS

This research was supported by a National Institutes of Health research project (no. 2022-ER1102-00), South Korea.

TABLE II

BASELINE CHARACTERISTICS

| | Study dataset | | | Extra-validation dataset | | |
|---|---|---|---|---|---|---|
| | Total | Control | Case† | Total | Control | Case† |
| Number, N | 12,809 | 11,571 | 1,238 | 2,019 | 1,966 | 32 |
| Age, means (STD) | 62.5 (12.05) | 62.1 (12.15) | 66.0 (10.46) | 56.3 (11.94) | 56.2 (11.92) | 60.5 (12.11) |
| Male, N (%) | 6,530 (50.98) | 5,859 (50.64) | 671 (54.20) | 1,094 (54.19) | 1,058 (53.81) | 36 (67.92) |
| Female, N (%) | 6,279 (49.02) | 5,712 (49.36) | 567 (45.80) | 925 (45.81) | 908 (46.19) | 17 (32.08) |
| Metformin, N (%) | 4,709 (36.76) | 4,263 (36.84) | 446 (36.03) | 734 (36.35) | 719 (36.57) | 15 (28.30) |
| Sulfonylurea, N (%) | 2,017 (15.75) | 1,848 (15.97) | 169 (13.65) | 172 (8.52) | 168 (8.55) | 4 (7.55) |
| Dipeptidyl Peptidase-4 inhibitor, N (%) | 1,713 (13.37) | 1,540 (13.31) | 173 (13.97) | 64 (3.17) | 63 (3.20) | 1 (1.89) |
| Meglitinide, N (%) | 90 (0.70) | 86 (0.74) | 4 (0.32) | 34 (1.68) | 32 (1.63) | 2 (3.77) |
| Thiazolidinedione, N (%) | 190 (1.48) | 179 (1.55) | 11 (0.89) | 2 (0.10) | 2 (0.10) | N/A |
| A-glucosidase inhibitor, N (%) | 105 (0.82) | 103 (0.89) | 2 (0.16) | 28 (1.39) | 28 (1.42) | N/A |
| Insulin, N (%) | 2,017 (15.75) | 1,848 (15.97) | 169 (13.65) | N/A | N/A | N/A |
| Glucagon like peptide-1 agonist, N (%) | 3 (0.02) | 3 (0.03) | N/A | N/A | N/A | N/A |
| Sodium Glucose Co-Transporter 2 inhibitor, N (%) | 104 (0.81) | 86 (0.74) | 18 (1.45) | N/A | N/A | N/A |
| Angiotensin II receptor blocker, N (%) | 3,391 (26.47) | 3,036 (26.24) | 355 (28.68) | 92 (4.56) | 88 (4.48) | 4 (7.55) |
| Angiotensin converting enzyme inhibitor, N (%) | 137 (1.07) | 118 (1.02) | 19 (1.53) | 19 (0.94) | 18 (0.92) | 1 (1.89) |
| Calcium-channel blocker, N (%) | 3,299 (25.76) | 2,982 (25.77) | 317 (25.61) | 228 (11.29) | 224 (11.39) | 4 (7.55) |
| Diuretics, N (%) | 1,486 (11.60) | 1,348 (11.65) | 138 (11.15) | 41 (2.03) | 39 (1.98) | 2 (3.77) |
| Beta blocker, N (%) | 1,120 (8.74) | 881 (7.61) | 239 (19.31) | 1 (0.05) | 1 (0.05) | N/A |
| Statin, N (%) | 4,504 (35.16) | 3,897 (33.68) | 607 (49.03) | 478 (23.68) | 463 (23.55) | 15 (28.30) |
| Fibrate, N (%) | 111 (0.87) | 105 (0.91) | 6 (0.48) | 52 (2.58) | 50 (2.54) | 2 (3.77) |
| Ezetimibe, N (%) | 231 (1.80) | 174 (1.50) | 57 (4.60) | 35 (1.73) | 34 (1.73) | 1 (1.89) |
| Omega-3, N (%) | 173 (1.35) | 153 (1.32) | 20 (1.62) | 18 (0.89) | 18 (0.92) | N/A |
| ETC dyslipidemia, N (%) | 6 (0.05) | 5 (0.04) | 1 (0.08) | 13 (0.64) | 13 (0.66) | N/A |
| Aspirin, N (%) | 1,756 (13.71) | 1,564 (13.52) | 192 (15.51) | 288 (14.26) | 279 (14.19) | 9 (16.98) |
| Clopidogrel, N (%) | 1,382 (10.79) | 1,107 (9.57) | 275 (22.21) | 70 (3.47) | 67 (3.41) | 3 (5.66) |
| Cilostazol, N (%) | 796 (6.21) | 730 (6.31) | 66 (5.33) | 60 (2.97) | 58 (2.95) | 2 (3.77) |
| Glycoprotein IIb/IIIa antagonist, N (%) | 19 (0.15) | 19 (0.16) | N/A | N/A | N/A | N/A |
| ETC Anti-platelet, N (%) | 796 (6.21) | 687 (5.94) | 109 (8.80) | N/A | N/A | N/A |

continued

(CONTINUED) TABLE II

BASELINE CHARACTERISTICS

| | Study dataset | | | Extra-validation dataset | | |
|---|---|---|---|---|---|---|
| | Total | Control | Case† | Total | Control | Case† |
| Body mass index, means (STD) | 24.9 (3.75) | 24.9 (3.74) | 25.4 (3.78) | 25.2 (3.62) | 25.2 (3.64) | 24.5 (3.11) |
| Systolic blood pressure, means (STD) | 124.1 (14.30) | 124.2 (14.24) | 123.4 (14.82) | 128.0 (13.36) | 128.0 (13.27) | 126.1 (16.30) |
| Diastolic blood pressure, means (STD) | 75.7 (8.96) | 75.8 (8.91) | 74.3 (9.34) | 76.4 (8.54) | 76.4 (8.53) | 75.7 (9.13) |
| Pulse rate, means (STD) | 75.6 (10.62) | 75.9 (10.54) | 73.6 (11.13) | 79.9 (10.63) | 79.9 (10.59) | 78.7 (12.01) |
| HbA1c, means (STD) | 6.9 (1.01) | 6.9 (1.00) | 6.9 (1.03) | 7.2 (1.09) | 7.2 (1.08) | 7.5 (1.09) |
| Blood glucose, means (STD) | 151.3 (46.35) | 151.6 (45.87) | 148.9 (50.49) | 146.7 (46.10) | 146.5 (46.21) | 155.9 (41.31) |
| Total cholesterol, means (STD) | 160.7 (32.86) | 160.7 (32.42) | 160.4 (36.70) | 166.2 (32.09) | 165.9 (31.84) | 175.0 (39.65) |
| Triglyceride, means (STD) | 143.8 (69.12) | 143.3 (68.51) | 148.7 (74.47) | 147.2 (66.68) | 146.6 (66.35) | 169.5 (75.41) |
| High-density lipoprotein, means (STD) | 47.4 (12.37) | 47.6 (12.30) | 45.5 (12.83) | 46.4 (10.82) | 46.5 (10.83) | 42.7 (9.85) |
| Low-density lipoprotein, means (STD) | 93.0 (26.84) | 92.9 (26.48) | 93.8 (29.93) | 94.6 (29.56) | 94.5 (29.47) | 99.5 (32.63) |
| Creatinine, means (STD) | 0.9 (0.45) | 0.9 (0.44) | 0.9 (0.47) | 1.0 (0.54) | 1.0 (0.53) | 1.2 (0.78) |
| Aspartate aminotransferase, means (STD) | 26.7 (13.42) | 26.7 (13.55) | 26.9 (12.11) | 24.5 (8.36) | 24.5 (8.36) | 24.7 (8.38) |
| Alanine aminotransferase, means (STD) | 24.0 (13.62) | 23.9 (13.38) | 25.1 (15.66) | 25.5 (12.83) | 25.4 (12.77) | 27.0 (15.14) |
| Gamma(γ)-glutamyl transferase, means (STD) | 39.5 (37.36) | 39.2 (37.41) | 41.9 (36.79) | 37.3 (30.38) | 37.2 (30.39) | 42.1 (29.68) |
| Alkaline phosphatase, means (STD) | 79.8 (25.50) | 80.1 (25.52) | 77.0 (25.20) | 165.8 (79.31) | 166.2 (79.17) | 152.2 (84.11) |
| Hypertension, N (%) | 6,995 (54.61) | 6,188 (53.48) | 807 (65.19) | 894 (44.28) | 867 (44.10) | 27 (50.94) |
| Dyslipidemia, N (%) | 5,435 (42.43) | 4,748 (41.03) | 687 (55.49) | 790 (39.13) | 767 (39.01) | 23 (43.40) |
| Stroke, N (%) | 18 (0.14) | 18 (0.16) | N/A | 135 (6.69) | 130 (6.61) | 5 (9.43) |
| Cerebrovascular diseases, N (%) | 1,317 (10.28) | 1,212 (10.47) | 105 (8.48) | 135 (6.69) | 130 (6.61) | 5 (9.43) |
| Dementia, N (%) | 207 (1.62) | 190 (1.64) | 17 (1.37) | 18 (0.89) | 18 (0.92) | 0 (0.00) |
| Neuropathy, N (%) | 706 (5.51) | 661 (5.71) | 45 (3.63) | 283 (14.02) | 270 (13.73) | N/A |
| Retinopathy, N (%) | 191 (1.49) | 179 (1.55) | 12 (0.97) | 220 (10.90) | 215 (10.94) | 5 (9.43) |
| Chronic kidney disease, N (%) | 356 (2.78) | 319 (2.76) | 37 (2.99) | 263 (13.03) | 252 (12.82) | 11 (20.75) |
| Parkinson's disease, N (%) | 96 (0.75) | 81 (0.70) | 15 (1.21) | 5 (0.25) | 4 (0.20) | 1 (1.89) |
| Peripheral vascular disease, N (%) | 150 (1.17) | 136 (1.18) | 14 (1.13) | 133 (6.59) | 125 (6.36) | 8 (15.09) |
| Lower limb amputation, N (%) | 3 (0.02) | 3 (0.03) | N/A | N/A | N/A | N/A |
| Cancer, N (%) | 685 (5.35) | 649 (5.61) | 36 (2.91) | 103 (5.10) | 100 (5.09) | 3 (5.66) |
| Proliferative diabetic retinopathy, N (%) | 247 (1.93) | 231 (2.00) | 16 (1.29) | 53 (2.63) | 53 (2.70) | N/A |

TABLE III

STUDY DATA RESULTS

| | AUC | Accuracy | Sensitivity | Specificity | Balanced accuracy |
|---|---|---|---|---|---|
| XGB | 84.84 (83.49-86.19) | 76.03 (74.68-77.39) | 76.01 (74.70-77.32) | 76.03 (74.67-77.40) | 76.02 (74.69-77.35) |
| RF | 83.46 (82.18-84.73) | 75.12 (73.61-76.63) | 75.04 (73.43-76.65) | 75.13 (73.63-76.63) | 75.08 (73.53-76.64) |
| AB | 79.57 (77.96-81.19) | 72.53 (70.80-74.25) | 72.37 (70.65-74.09) | 72.54 (70.82-74.27) | 72.46 (70.74-74.18) |
| LGBM | 83.81 (82.26-85.36) | 76.18 (74.41-77.95) | 76.09 (74.52-77.66) | 76.19 (74.40-77.98) | 76.14 (74.47-77.81) |

TABLE IV

EXTERNAL VALIDATION DATA RESULTS

| | AUC | Accuracy | Sensitivity | Specificity | Balanced accuracy |
|---|---|---|---|---|---|
| XGB | 69.81 | 66.37 | 66.04 | 66.38 | 66.21 |
| RF | 72.76 | 68.1 | 67.92 | 68.11 | 68.02 |
| AB | 68.82 | 65.82 | 66.04 | 65.82 | 65.93 |
| LGBM | 70.97 | 63.25 | 64.15 | 63.22 | 63.69 |

## REFERENCES


[1]. Noble, W.S., *Support vector machine applications in computational biology*. Kernel methods in computational biology, 2004. **71**: p. 92.

[2]. Aruna, S. and S. Rajagopalan, *A novel SVM based CSSFFS feature selection algorithm for detecting breast cancer*. International journal of computer applications, 2011. **31**(8): p. 14-20.

[3]. Lakhani, P. and B. Sundaram, *Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks.* Radiology, 2017. **284**(2): p. 574-582.

[4]. Yasaka, K., et al., *Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study.* Radiology, 2018. **286**(3): p. 887-896.

[5]. Christ, P.F., et al. *Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields*. in *International conference on medical image computing and computer-assisted intervention*. 2016. Springer.

[6]. Wong, J., et al., *Using machine learning to identify health outcomes from electronic health record data.* Curr Epidemiol Rep, 2018. **5**(4): p. 331-342.

[7]. Goldstein, B.A., et al., *Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review.* J Am Med Inform Assoc, 2017. **24**(1): p. 198-208.

[8]. Goldstein, B.A., A.M. Navar, and R.E. Carter, *Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges.* Eur Heart J, 2017. **38**(23): p. 1805-1814.

[9]. Hernesniemi, J.A., et al., *Extensive phenotype data and machine learning in prediction of mortality in acute coronary syndrome - the MADDEC study.* Ann Med, 2019. **51**(2): p. 156-163.

[10]. Collins, G.S., et al., *Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement.* Bmj, 2015. **350**: p. g7594.

[11]. Livne, M., et al., *Boosted Tree Model Reforms Multimodal Magnetic Resonance Imaging Infarct Prediction in Acute Stroke.* Stroke, 2018. **49**(4): p. 912-918.

[12]. Zhang, Z., K.M. Ho, and Y. Hong, *Machine learning for the prediction of volume responsiveness in patients with oliguric acute kidney injury in critical care.* Crit Care, 2019. **23**(1): p. 112.

[13]. Rigatti, S.J., *Random Forest.* J Insur Med, 2017. **47**(1): p. 31-39.

[14]. Collins, M., R.E. Schapire, and Y. Singer, *Logistic regression, AdaBoost and Bregman distances.* Machine Learning, 2002. **48**(1-3): p. 253-285.

[15]. Ke, G., et al., *Lightgbm: A highly efficient gradient boosting decision tree.* Advances in neural information processing systems, 2017. **30**.

# Long-term trends in the prevalence of rheumatoid arthritis and osteoarthritis in South Korea, from 1998 to 2021: a 24-year nationally representative serial study

Jaeyu Park,[1] Hyeri Lee, [1] Minji Kim, [1] Wonyoung Cho,[1*]
[1]*Center for Digital Health, Medical Science Research Institute, Kyung Hee University Medical Center, Kyung Hee University College of Medicine, Seoul, South Korea*
**Correspondence: Wonyoung Cho (bourbaki10@gmail.)*

***Abstract*— Studies on the trends in the prevalence of rheumatoid arthritis (RA) and osteoarthritis (OA) are limited, particularly during the COVID-19 pandemic. This study aimed to analyze the temporal trend of RA and OA in Korean adults from 1998 to 2021, including the COVID-19 pandemic period. The Korea National Health and Nutrition Examination Survey (KNHANES) data on Korean adults aged ≥19 years were analyzed to investigate the prevalence of RA and OA from 1998 to 2021. The prevalence trends were compared across subgroups, and $\beta_{diff}$ ($\beta$ difference) was calculated. Odds ratios (ORs) were computed for each disease to examine changes in disease prevalence before and during the pandemic in order to determine the impact of the pandemic on disease prevalence. Among 163,221 Korean adults, the prevalence of RA and OA showed a steady decrease from 1998 to 2019 but increased after the onset of the COVID-19 pandemic. Vulnerable groups, including participants aged ≥60 years, urban, and participants with higher education level showed higher prevalence of OA, whereas no particularly vulnerable population was observed for RA. Our findings provide an insight into the long-term trends of RA and OA in Korean adults and highlight a novel perspective on the impact of COVID-19 on disease.**

## I. Introduction

The estimated global prevalence rate of rheumatoid arthritis (RA) is approximately 0.24%.[1] Previous studies have demonstrated that this rate has remained relatively consistent over time.[1] However, in South Korea, studies have reported a decrease in the prevalence rate of RA before the COVID-19 pandemic, followed by an increase during the pandemic.[2] Similar trends have been observed in the prevalence rate of osteoarthritis (OA) globally, with an increasing trend; however, in South Korea, it decreased before the pandemic and increased during the pandemic.[3,4] The implementation of social distancing measures during the COVID-19 pandemic may have directly or indirectly influenced the prevalence rates of RA and OA.[2] Moreover, changes in physical activity, dietary habits, and mental health due to the COVID-19 restrictions may have influenced these prevalence rates.[5] However, further analysis and research are needed to determine the direct relationship between RA, OA, and the COVID-19 pandemic. Furthermore, some medications used to treat RA, such as corticosteroids and disease-modifying antirheumatic drugs, may suppress the immune system and increase the risk of infection,[6] thereby, making the patients more vulnerable to COVID-19 and its complications.[6] Therefore, it is crucial for patients with RA and COVID-19 to prioritize their health management.[7] This study aimed to investigate the long-term trends in the prevalence of RA and OA using a representative large-scale dataset from 1998 to 2021. Additionally, we analyzed the specific social strata that were vulnerable during the COVID-19 period. Furthermore, our research findings may provide valuable information for developing policies for the treatment and prevention of these conditions following the COVID-19 pandemic.

## II. Methods

### A. *Patient selection and data collection*

This study utilized data from the Korea National Health and Nutrition Examination Survey (KNHANES) conducted between 1998 and 2021 by the Korea Centers for Disease Control and Prevention Agency (KDCA).8,9 The study population included adults aged ≥19 years, and the collected data included information on age, sex, residence, body mass index (BMI), education level, income, alcohol consumption, smoking status, and history of RA and OA. A nationally representative sample of 163,221 participants was used to investigate the prevalence of RA and OA before and during the COVID-19 pandemic. The survey was conducted over 24 years and the number of participants surveyed in each year group was as follows: 75,691 in 1998–2005; 16,695 in 2007–2009; 18,126 in 2010–2012; 16,754 in 2013–2015; 24,514 in 2016–2019; 5,839 in 2020; and 5,602 in 2021.

The study protocol has been endorsed by Kyung Hee University.

### B. *Main outcome*

The objective of our study was to investigate the risk factors associated with two most common types of arthritis, RA and OA, over a period of 24 years from 1998 to 2021. To achieve our research objective, we surveyed a large sample of participants and asked them the question: “Have you ever been diagnosed with RA or OA by a doctor?” Based on their answers, we categorized the participants into three groups: RA, OA, and both.[10] We collected data on various potential risk factors associated with the development of RA and OA,

such as age, sex, lifestyle habits, and socioeconomic status. We conducted statistical analyses to examine the associations between these risk factors and the development of RA and OA and to identify any patterns or trends that emerged over 24 years.

### C. *Covariates*

Covariates included age (19–29, 30–39, 40–49, 50–59, 60–69, 70–79, and ≥80 years), sex, region of residence (urban and rural)[11], BMI group, household income (lowest, second, third, and highest quartile), education level (elementary school or lower, middle school, high school, and college or higher education), alcohol consumption (1–5 days/month, ≥6 days/month, and non-drinker), and smoking status (non-smoker, ex-smoker, and smoker). BMI was categorized into underweight (<18.5 kg/m$^2$), normal weight (18.5-22.9 kg/m$^2$), overweight (23–25 kg/m$^2$), and obese (≥25.0 kg/m$^2$) according to the Asian-Pacific guidelines.[12,13]

### D. *Statistical analyses*

The results of this study were presented using qualitative data expressed as proportions or percentages. Weighted multivariate regression model analyses were conducted to compare the estimates of each related factor before and after the COVID-19 pandemic, with weighted odds ratios (ORs) with 95% confidence intervals (CIs).[14] The prevalence of RA and OA was calculated using data from the KNHANES, spanning from 1998 to 2021, stratified by year group. Weighted complex sampling analysis was performed to ensure accurate estimation. Binomial or linear logistic regression models were used to compute the ORs with 95% CIs or β-coefficients with 95% CIs. To ensure robustness of the main findings, a stratification analysis was performed using variables such as sex, educational level, region of residence, and income in all the regression models. Furthermore, the ratio of ORs was calculated to estimate the interaction term of each risk factor and identify groups that were more vulnerable to the prevalence of diagnosed RA and OA during the pandemic. Overall, this study aimed to provide a comprehensive analysis of the impact of the COVID-19 pandemic on the prevalence of RA and OA and to identify the factors that contribute to vulnerability to these conditions. The SAS software (version 9.4; SAS Institute, Cary, NC, USA) was used for statistical analyses, with a two-sided test, and a p-value ≤0.05 was considered statistically significant.[14]

## III. Results

A statistically significant decrease in the weighted prevalence of RA was observed pre-pandemic (1998–2005, 1.84% [95% CI: 1.64 to 2.03]; 2007–2009, 1.88% [95% CI: 1.65 to 2.10]; 2010–2012, 1.49% [95% CI: 1.30 to 1.68]; 2013–2015, 1.46% [95% CI: 1.27 to 1.64]; and 2016–2019, 1.58% [95% CI: 1.42 to 1.75]), whereas a slight increase was observed during the pandemic (2020, 1.23% [95% CI: 0.92 to 1.54] and 2021, 1.36% [95% CI: 1.02 to 1.69]). Similarly, a statistical significant decrease in the weighted prevalence of OA was observed, pre-pandemic (1998–2005, 10.49% [95% CI: 10.12 to 10.86]; 2007–2009, 8.76% [95% CI: 8.24 to 9.27]; 2010–2012, 7.91% [95% CI: 7.44 to 8.37]; 2013–2015, 8.08% [95% CI: 7.60 to 8.57]; 2016–2019, 8.38% [95% CI: 7.95 to 8.82]), whereas a slight increase was observed during the pandemic (2020, 8.04% [95% CI: 7.18 to 8.89] and 2021, 8.27% [95% CI: 7.35 to 9.19]). Similar patterns and trends were observed for the prevalence of OA and RA stratified by age, sex, region of residence, education level, and household income.

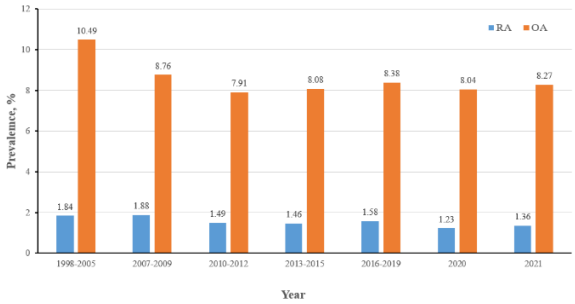



Fig. 1 Twenty-four-year trends in the prevalence of rheumatoid arthritis and osteoarthritis

Table 3 shows the pandemic-related effects on vulnerable groups. A statistically significant difference was observed among vulnerable individuals with OA as follows: participants ≥60 years old (ratio of ORs, 1.222 [95% CI: 1.011 to 1.477]), urban residents (ratio of ORs, 1.289 [95% CI: 1.007 to 1.650]), and participants with a high level of education (ratio of ORs, 1.360 [95% CI: 1.119 to 1.653]). In contrast, no significant difference in pandemic-related effects was observed among vulnerable individuals with RA.

## IV. Conclusions

Our study identified long-term trends in the prevalence of RA and OA over a 24-year period from 1998 to 2021, with a particular focus on the impact of the COVID-19 pandemic. The results showed a consistent decline in the prevalence of both RA and OA until the year 2020, followed by an increase in 2021, Notably, OA exhibited a higher prevalence among specific vulnerable groups, such as individuals over 60 years of age, urban residents, and those with a high education level, whereas RA did not show a particularly

vulnerable population. It would be beneficial for government policy researchers to devise personalized policies targeted at the vulnerable groups affected by OA. While our study did not find any evidence of a relationship between the COVID-19 pandemic and the prevalence of RA has yet been identified, additional follow-up studies based on our study findings would be helpful in further exploring this topic.

## V. ACKNOWLEDGMENTS

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number : HV22C0233).

Table 1. Baseline characteristics.

| | Total | 1998–2005 | 2007–2009 | 2010–2012 | 2013–2015 | 2016–2019 | 2020 | 2021 |
|---|---|---|---|---|---|---|---|---|
| Overall, n | 163,221 | 75,691 | 16,695 | 18,126 | 16,754 | 24,514 | 5,839 | 5,602 |
| Age (years), weighted % (95% CI) | | | | | | | | |
| 19–29 | 18.18 (17.82 to 18.53) | 20.81 (20.26 to 21.35) | 19.26 (18.15 to 20.36) | 17.94 (16.87 to 19.00) | 17.00 (16.09 to 17.91) | 16.52 (15.71 to 17.33) | 16.19 (14.65 to 17.72) | 15.91 (14.10 to 17.73) |
| 30–39 | 20.88 (20.47 to 21.29) | 25.02 (24.32 to 25.73) | 22.66 (21.44 to 23.88) | 21.09 (20.01 to 22.17) | 19.53 (18.41 to 20.65) | 17.97 (17.03 to 18.90) | 16.77 (14.85 to 18.69) | 16.32 (14.61 to 18.04) |
| 40–49 | 21.70 (21.35 to 22.06) | 23.35 (22.76 to 23.93) | 22.76 (21.72 to 23.80) | 22.19 (21.17 to 23.21) | 21.29 (20.37 to 22.21) | 20.41 (19.61 to 21.20) | 19.35 (17.61 to 21.09) | 18.99 (17.27 to 20.70) |
| 50–59 | 17.85 (17.56 to 18.14) | 14.22 (13.82 to 14.62) | 16.53 (15.76 to 17.31) | 18.31 (17.51 to 19.12) | 19.81 (18.98 to 20.64) | 20.14 (19.45 to 20.82) | 20.03 (18.60 to 21.46) | 19.78 (18.25 to 21.31) |
| 60–69 | 11.88 (11.64 to 12.12) | 10.71 (10.34 to 11.08) | 10.35 (9.78 to 10.91) | 10.67 (10.10 to 11.23) | 11.53 (10.91 to 12.15) | 13.37 (12.75 to 14.00) | 15.04 (13.65 to 16.43) | 16.06 (14.75 to 17.38) |
| 70–79 | 7.38 (7.19 to 7.58) | 5.60 (5.32 to 5.88) | 6.66 (6.17 to 7.14) | 7.70 (7.16 to 8.23) | 8.23 (7.70 to 8.77) | 8.33 (7.85 to 8.81) | 8.85 (7.71 to 9.98) | 8.93 (7.80 to 10.06) |
| ≥80 | 2.13 (2.03 to 2.23) | 0.29 (0.24 to 0.35) | 1.78 (1.56 to 2.00) | 2.11 (1.86 to 2.36) | 2.61 (2.34 to 2.88) | 3.27 (2.98 to 3.56) | 3.78 (3.05 to 4.50) | 4.01 (3.28 to 4.74) |
| Sex, weighted % (95% CI) | | | | | | | | |
| Male | 49.39 (49.14 to 49.64) | 48.63 (48.32 to 48.94) | 49.57 (48.85 to 50.29) | 49.55 (48.83 to 50.28) | 49.50 (48.79 to 50.22) | 49.78 (49.17 to 50.40) | 49.84 (48.73 to 50.95) | 49.82 (48.52 to 51.12) |
| Female | 50.61 (50.36 to 50.86) | 51.37 (51.06 to 51.68) | 50.43 (49.71 to 51.15) | 50.45 (49.72 to 51.17) | 50.50 (49.78 to 51.21) | 50.22 (49.60 to 50.83) | 50.16 (49.05 to 51.27) | 50.18 (48.88 to 51.48) |
| Region of residence, weighted % (95% CI) | | | | | | | | |
| Urban | 82.15 (81.14 to 83.16) | 81.35 (80.57 to 82.14) | 80.57 (77.48 to 83.66) | 79.80 (76.52 to 83.08) | 82.12 (79.18 to 85.05) | 84.59 (82.14 to 87.04) | 84.57 (79.45 to 89.68) | 83.80 (78.76 to 88.84) |
| Rural | 17.85 (16.84 to 18.86) | 18.65 (17.86 to 19.43) | 19.43 (16.34 to 22.52) | 20.20 (16.92 to 23.48) | 17.88 (14.95 to 20.82) | 15.41 (12.96 to 17.86) | 15.43 (10.32 to 20.55) | 16.20 (11.16 to 21.24) |
| BMI group, weighted % (95% CI) | | | | | | | | |
| Underweight | 3.52 (3.39 to 3.65) | 1.02 (0.89 to 1.14) | 4.70 (4.28 to 5.13) | 4.81 (4.41 to 5.21) | 4.49 (4.10 to 4.88) | 3.93 (3.63 to 4.23) | 4.13 (3.45 to 4.81) | 4.17 (3.57 to 4.78) |
| Normal weight and overweight | 49.31 (48.82 to 49.81) | 14.58 (13.29 to 15.88) | 62.97 (62.04 to 63.89) | 62.42 (61.49 to 63.35) | 62.33 (61.46 to 63.20) | 60.64 (59.87 to 61.42) | 56.40 (54.78 to 58.03) | 57.52 (55.80 to 59.24) |
| Obesity | 27.05 (26.68 to 27.41) | 7.15 (6.48 to 7.82) | 31.79 (30.89 to 32.68) | 32.15 (31.21 to 33.10) | 32.92 (32.03 to 33.80) | 34.98 (34.20 to 35.76) | 38.15 (36.60 to 39.71) | 36.89 (35.11 to 38.68) |
| Unknown | 20.12 (19.52 to 20.72) | 77.25 (75.26 to 79.25) | 0.54 (0.38 to 0.70) | 0.62 (0.46 to 0.77) | 0.26 (0.17 to 0.35) | 0.45 (0.33 to 0.56) | 1.31 (0.98 to 1.65) | 1.41 (1.01 to 1.81) |

Table 2. Prevalence of RA.

| Year | Pre-pandemic | | | | | During the pandemic | | Trends in the pre-pandemic era, β (95% CI) | Trends in the pandemic era, β (95% CI) | Trends difference, $\beta_{diff}$ (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1998–2005* | 2007–2009 | 2010–2012 | 2013–2015 | 2016–2019 | 2020 | 2021 | | | |
| RA | | | | | | | | | | |
| Overall | 1.84 (1.64 to 2.03) | 1.88 (1.65 to 2.10) | 1.49 (1.30 to 1.68) | 1.46 (1.27 to 1.64) | 1.58 (1.42 to 1.75) | 1.23 (0.92 to 1.54) | 1.36 (1.02 to 1.69) | **-0.116 (-0.181 to -0.050)** | 0.129 (-0.333 to 0.591) | 0.244 (-0.222 to 0.711) |
| Age group | | | | | | | | | | |
| Age (19–60 years) | 1.38 (1.19 to 1.57) | 1.21 (1.00 to 1.42) | 1.03 (0.83 to 1.23) | 0.97 (0.77 to 1.17) | 0.88 (0.72 to 1.04) | 0.56 (0.27 to 0.85) | 0.73 (0.40 to 1.06) | **-0.147 (-0.208 to -0.087)** | 0.173 (-0.263 to 0.610) | 0.321 (-0.120 to 0.761) |
| Age (≥60 years) | 4.19 (3.54 to 4.85) | 4.76 (4.06 to 5.46) | 3.26 (2.79 to 3.73) | 3.15 (2.65 to 3.66) | 3.70 (3.24 to 4.16) | 2.98 (2.20 to 3.77) | 2.89 (2.10 to 3.68) | **-0.268 (-0.456 to -0.079)** | -0.092 (-1.197 to 1.013) | 0.176 (-0.945 to 1.296) |
| Sex | | | | | | | | | | |
| Male | 1.03 (0.82 to 1.23) | 0.98 (0.73 to 1.23) | 0.69 (0.46 to 0.91) | 0.80 (0.59 to 1.02) | 0.73 (0.55 to 0.91) | 0.53 (0.28 to 0.78) | 0.82 (0.46 to 1.19) | **-0.088 (-0.148 to -0.028)** | 0.296 (-0.148 to 0.739) | 0.383 (-0.064 to 0.831) |
| Female | 2.64 (2.33 to 2.95) | 2.76 (2.41 to 3.10) | 2.28 (1.96 to 2.59) | 2.10 (1.80 to 2.40) | 2.42 (2.16 to 2.69) | 1.92 (1.34 to 2.51) | 1.89 (1.41 to 2.36) | **-0.141 (-0.257 to -0.025)** | -0.038 (-0.787 to 0.711) | 0.103 (-0.655 to 0.861) |
| Region of residence | | | | | | | | | | |
| Urban | 1.65 (1.45 to 1.85) | 1.69 (1.45 to 1.94) | 1.37 (1.16 to 1.58) | 1.48 (1.27 to 1.69) | 1.50 (1.33 to 1.68) | 1.21 (0.86 to 1.56) | 1.22 (0.86 to 1.59) | **-0.079 (-0.153 to -0.005)** | 0.012 (-0.499 to 0.524) | 0.091 (-0.425 to 0.608) |
| Rural | 2.69 (2.13 to 3.25) | 2.64 (2.09 to 3.19) | 1.96 (1.47 to 2.44) | 1.34 (0.97 to 1.71) | 2.00 (1.57 to 2.42) | 1.32 (0.71 to 1.93) | 2.05 (1.16 to 2.94) | **-0.266 (-0.407 to -0.125)** | 0.725 (-0.343 to 1.794) | 0.991 (-0.086 to 2.069) |
| Education | | | | | | | | | | |
| High school or lower education | 2.48 (2.19 to 2.77) | 2.67 (2.34 to 3.00) | 2.21 (1.92 to 2.51) | 2.27 (1.95 to 2.58) | 2.68 (2.37 to 2.98) | 2.14 (1.57 to 2.71) | 2.14 (1.61 to 2.67) | -0.051 (-0.160 to 0.057) | 0.000 (-0.797 to 0.797) | 0.051 (-0.753 to 0.855) |
| College or higher education | 0.67 (0.48 to 0.85) | 0.66 (0.43 to 0.90) | 0.62 (0.39 to 0.84) | 0.81 (0.56 to 1.05) | 0.71 (0.54 to 0.87) | 0.57 (0.26 to 0.87) | 0.81 (0.38 to 1.25) | -0.011 (-0.080 to 0.057) | 0.246 (-0.284 to 0.777) | 0.258 (-0.277 to 0.792) |
| Income | | | | | | | | | | |
| Income (lowest-second quartile) | 2.69 (2.36 to 3.03) | 2.52 (2.16 to 2.88) | 1.83 (1.52 to 2.14) | 1.85 (1.53 to 2.18) | 2.26 (1.98 to 2.54) | 1.94 (1.34 to 2.54) | 1.82 (1.26 to 2.38) | -0.103 (-0.220 to 0.014) | -0.124 (-0.944 to 0.696) | -0.021 (-0.849 to 0.808) |
| Income (third-highest quartile) | 1.17 (0.97 to 1.36) | 1.42 (1.16 to 1.68) | 1.22 (0.99 to 1.46) | 1.19 (0.96 to 1.41) | 1.13 (0.94 to 1.32) | 0.81 (0.50 to 1.11) | 1.09 (0.71 to 1.46) | **-0.104 (-0.173 to -0.035)** | 0.281 (-0.205 to 0.768) | 0.385 (-0.106 to 0.877) |

Table 3. Prevalence of OA.

| Year | Pre-pandemic | | | | | During the pandemic | | Trends in the pre-pandemic era, β (95% CI) | Trends in the pandemic era, β (95% CI) | Trends difference, $\beta_{diff}$ (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1998–2005* | 2007–2009 | 2010–2012 | 2013–2015 | 2016–2019 | 2020 | 2021 | | | |
| OA | | | | | | | | | | |
| Overall | 10.49 (10.12 to 10.86) | 8.76 (8.24 to 9.27) | 7.91 (7.44 to 8.37) | 8.08 (7.60 to 8.57) | 8.38 (7.95 to 8.82) | 8.04 (7.18 to 8.89) | 8.27 (7.35 to 9.19) | **-0.331 (-0.510 to -0.152)** | 0.231 (-1.087 to 1.549) | 0.562 (-0.768 to 1.892) |
| Age group | | | | | | | | | | |
| Age (19–60 years) | 5.62 (5.35 to 5.90) | 4.57 (4.15 to 4.99) | 3.35 (3.00 to 3.71) | 3.20 (2.85 to 3.55) | 3.01 (2.71 to 3.30) | 2.35 (1.80 to 2.90) | 2.79 (2.19 to 3.39) | **-0.601 (-0.718 to -0.484)** | 0.438 (-0.366 to 1.242) | **1.040 (0.227 to 1.852)** |
| Age (≥60 years) | 34.95 (33.78 to 36.13) | 26.85 (25.37 to 28.33) | 25.59 (24.28 to 26.90) | 25.04 (23.66 to 26.43) | 24.54 (23.50 to 25.58) | 22.91 (20.86 to 24.97) | 21.69 (19.53 to 23.85) | **-1.620 (-2.102 to -1.138)** | -1.223 (-4.281 to 1.835) | 0.397 (-2.698 to 3.493) |
| Sex | | | | | | | | | | |
| Male | 5.18 (4.86 to 5.49) | 4.28 (3.77 to 4.79) | 2.92 (2.50 to 3.33) | 3.44 (2.99 to 3.89) | 3.19 (2.85 to 3.53) | 3.63 (2.86 to 4.40) | 4.07 (3.22 to 4.92) | **-0.257 (-0.417 to -0.098)** | 0.442 (-0.696 to 1.580) | 0.699 (-0.449 to 1.848) |
| Female | 15.52 (14.95 to 16.09) | 13.15 (12.32 to 13.98) | 12.81 (12.02 to 13.59) | 12.63 (11.87 to 13.40) | 13.54 (12.82 to 14.25) | 12.42 (10.99 to 13.85) | 12.44 (11.01 to 13.86) | **-0.391 (-0.687 to -0.096)** | 0.017 (-2.133 to 2.168) | 0.409 (-1.762 to 2.579) |
| Region of residence | | | | | | | | | | |
| Urban | 8.59 (8.23 to 8.95) | 8.05 (7.50 to 8.59) | 6.97 (6.48 to 7.45) | 7.51 (7.00 to 8.03) | 7.79 (7.32 to 8.25) | 7.81 (6.89 to 8.74) | 7.58 (6.65 to 8.52) | -0.135 (-0.325 to 0.055) | -0.232 (-1.572 to 1.108) | -0.097 (-1.450 to 1.257) |
| Rural | 18.78 (17.50 to 20.06) | 11.70 (10.14 to 13.25) | 11.62 (10.25 to 12.99) | 10.70 (9.09 to 12.32) | 11.67 (10.27 to 13.07) | 9.26 (6.58 to 11.94) | 11.81 (9.18 to 14.44) | **-1.050 (-1.585 to -0.516)** | 2.556 (-1.329 to 6.441) | 3.606 (-0.316 to 7.528) |
| Education | | | | | | | | | | |
| High school or lower education | 14.56 (14.05 to 15.06) | 13.22 (12.48 to 13.97) | 12.96 (12.22 to 13.70) | 15.04 (14.18 to 15.91) | 15.87 (15.12 to 16.62) | 15.87 (14.32 to 17.41) | 14.75 (13.08 to 16.42) | **0.429 (0.145 to 0.714)** | -1.116 (-3.428 to 1.197) | -1.545 (-3.875 to 0.785) |
| College or higher education | 2.19 (1.95 to 2.44) | 1.92 (1.54 to 2.30) | 1.59 (1.25 to 1.93) | 1.79 (1.46 to 2.12) | 2.16 (1.88 to 2.44) | 2.42 (1.76 to 3.08) | 3.36 (2.65 to 4.06) | 0.096 (-0.042 to 0.234) | 0.932 (-0.042 to 1.905) | 0.835 (-0.148 to 1.819) |
| Income | | | | | | | | | | |
| Income (lowest-second quartile) | 15.75 (15.12 to 16.37) | 13.24 (12.39 to 14.09) | 11.91 (11.12 to 12.71) | 12.82 (11.86 to 13.77) | 13.62 (12.85 to 14.40) | 13.31 (11.74 to 14.88) | 14.23 (12.41 to 16.05) | -0.240 (-0.559 to 0.079) | 0.919 (-1.591 to 3.429) | 1.159 (-1.371 to 3.690) |
| Income (third-highest quartile) | 6.10 (5.78 to 6.42) | 5.59 (5.08 to 6.09) | 4.80 (4.34 to 5.26) | 4.87 (4.40 to 5.33) | 4.90 (4.51 to 5.30) | 4.92 (4.12 to 5.71) | 4.78 (4.00 to 5.56) | **-0.224 (-0.386 to -0.061)** | -0.134 (-1.221 to 0.953) | 0.090 (-1.009 to 1.189) |

Table 4. Risk factor of RA&OA.

| Variables | | Overall (2005–2021) | | Pre-COVID-19 pandemic (2005–2019) | | COVID-19 pandemic (2020–2021) | | Ratio of ORs (95% CI), pre-pandemic (reference) versus pandemic | p-value |
|---|---|---|---|---|---|---|---|---|---|
| | | Weighted OR (95% CI) | p-value | Weighted OR (95% CI) | p-value | Weighted OR (95% CI) | p-value | | |
| RA | | | | | | | | | |
| Age (years) | 19-60 | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | |
| | ≥60 | **3.596 (3.248 to 3.980)** | **<.0001** | **3.553 (3.197 to 3.948)** | **<.0001** | **4.665 (3.145 to 6.919)** | **<.0001** | 1.313 (0.873 to 1.975) | 0.191 |
| Sex | Female | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | |
| | Male | **0.335 (0.298 to 0.377)** | **<.0001** | **0.334 (0.295 to 0.379)** | **<.0001** | **0.348 (0.244 to 0.497)** | **<.0001** | 1.042 (0.715 to 1.519) | 0.831 |
| Region | Urban | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | |
| | Rural | **1.433 (1.272 to 1.615)** | **<.0001** | **1.424 (1.257 to 1.614)** | **<.0001** | **1.502 (1.015 to 2.223)** | **<.0001** | 1.055 (0.699 to 1.592) | 0.799 |
| Education | High school or lower education | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | |
| | College or higher education | **0.245 (0.214 to 0.280)** | **<.0001** | **0.241 (0.209 to 0.278)** | **<.0001** | **0.286 (0.192 to 0.427)** | **<.0001** | 1.187 (0.777 to 1.814) | 0.428 |
| Income | High | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | |
| | Low | **1.885 (1.710 to 2.079)** | **<.0001** | **1.864 (1.683 to 2.066)** | **<.0001** | **2.004 (1.460 to 2.752)** | **<.0001** | 1.075 (0.770 to 1.500) | 0.670 |
| OA | | | | | | | | | |
| Age (years) | 19-60 | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | |
| | ≥60 | **8.852 (8.467 to 9.254)** | **<.0001** | **8.839 (8.442 to 9.254)** | **<.0001** | **10.799 (8.986 to 12.980)** | **<.0001** | **1.222 (1.011 to 1.477)** | **0.038** |
| Sex | Female | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | |
| | Male | **0.256 (0.244 to 0.268)** | **<.0001** | **0.253 (0.241 to 0.265)** | **<.0001** | **0.286 (0.243 to 0.337)** | **<.0001** | 1.130 (0.953 to 1.340) | 0.159 |
| Region | Rural | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | |
| | Urban | **0.535 (0.501 to 0.571)** | **<.0001** | **0.523 (0.489 to 0.560)** | **<.0001** | **0.674 (0.531 to 0.854)** | **<.0001** | **1.289 (1.007 to 1.650)** | **0.044** |
| Education | High school or lower education | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | |
| | College or higher education | **0.117 (0.110 to 0.125)** | **<.0001** | **0.111 (0.103 to 0.119)** | **<.0001** | **0.151 (0.126 to 0.181)** | **<.0001** | **1.360 (1.119 to 1.653)** | **0.002** |
| Income | High | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | | 1.000 (reference) | |
| | Low | **2.864 (2.742 to 2.992)** | **<.0001** | **2.839 (2.712 to 2.971)** | **<.0001** | **3.129 (2.689 to 3.641)** | **<.0001** | 1.102 (0.941 to 1.291) | 0.229 |

## REFERENCES


1. Cross M, Smith E, Hoy D, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014; **73**(7): 1316-22.
2. Ahn SM, Eun S, Ji S, et al. Incidence of rheumatic diseases during the COVID-19 pandemic in South Korea. *Korean J Intern Med* 2023; **38**(2): 248-53.
3. Hamood R, Tirosh M, Fallach N, Chodick G, Eisenberg E, Lubovsky O. Prevalence and Incidence of Osteoarthritis: A Population-Based Retrospective Cohort Study. *J Clin Med* 2021; **10**(18).
4. Kang Y, Ahmad S. The emerging epidemics in recent: mpox. *Life Cycle* 2023; **3**: e3.
5. Schäfer C, Keyßer G. Lifestyle Factors and Their Influence on Rheumatoid Arthritis: A Narrative Review. *J Clin Med* 2022; **11**(23).
6. Shin YH, Shin JI, Moon SY, et al. Autoimmune inflammatory rheumatic diseases and COVID-19 outcomes in South Korea: a nationwide cohort study. *Lancet Rheumatol* 2021; **3**(10): e698-e706.
7. Machado PM, Verschueren P, Grainger R, et al. Impact of COVID-19 pandemic on the management of patients with RA: a survey of rheumatologists in six European countries. *Rheumatol Adv Pract* 2023; **7**(1): rkac108.
8. Koh HY, Kim TH, Sheen YH, et al. Serum heavy metal levels are associated with asthma, allergic rhinitis, atopic dermatitis, allergic multimorbidity, and airflow obstruction. *J Allergy Clin Immunol Pract* 2019; **7**(8): 2912-5.e2.
9. Kweon S, Kim Y, Jang MJ, et al. Data resource profile: the Korea National Health and Nutrition Examination Survey (KNHANES). *Int J Epidemiol* 2014; **43**(1): 69-77.
10. Hur N-W, Choi C-B, Uhm W-S, Bae S-C. The Prevalence and Trend of Arthritis in Korea: Results from Korea National Health and Nutrition Examination Surveys. *jkra* 2008; **15**(1): 11-26.
11. Yoo IK, Marshall DC, Cho JY, Yoo HW, Lee SW. N-Nitrosodimethylamine-contaminated ranitidine and risk of cancer in South Korea: a nationwide cohort study. *Life Cycle* 2021; **1**: e1.
12. Yoon SY, Park HW, Kim HJ, et al. National trends in the prevalence of chronic kidney disease among Korean adults, 2007-2020. *Sci Rep* 2023; **13**(1): 5831.
13. Eum S, Son JW, Min C, et al. Ethnic and sex differences in the distributions of body mass index and waist circumference among adults: a binationally representative study in South Korea and the United States. *Eur Rev Med Pharmacol Sci* 2023; **27**(5): 1889-903.
14. Lee SW. Regression analysis for continuous independent variables in medical research: statistical standard and guideline of Life Cycle Committee. *Life Cycle* 2022; **2**: e3.
15. Dewanjee S, Kandimalla R, Kalra RS, et al. COVID-19 and Rheumatoid Arthritis Crosstalk: Emerging Association, Therapeutic Options and Challenges. *Cells* 2021; **10**(12).
16. Sung YK, Cho SK, Choi CB, Bae SC. Prevalence and incidence of rheumatoid arthritis in South Korea. *Rheumatol Int* 2013; **33**(6): 1525-32.
17. Myasoedova E, Davis J, Matteson EL, Crowson CS. Is the epidemiology of rheumatoid arthritis changing? Results from a population-based incidence study, 1985-2014. *Ann Rheum Dis* 2020; **79**(4): 440-4.
18. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. *Arthritis research & therapy* 2002; **4**: 1-8.
19. Lee KJ, Seon SY, Noh B, An KO. Physical fitness changes in adolescents due to social distancing during the coronavirus disease pandemic in Korea. *PeerJ* 2022; **10**: e14494.
20. Di Giuseppe D, Bottai M, Askling J, Wolk A. Physical activity and risk of rheumatoid arthritis in women: a population-based prospective study. *Arthritis Res Ther* 2015; **17**(1): 40.
21. Heidari B. Knee osteoarthritis prevalence, risk factors, pathogenesis and features: Part I. *Caspian J Intern Med* 2011; **2**(2): 205-12.
22. Deane KD, Demoruelle MK, Kelmenson LB, Kuhn KA, Norris JM, Holers VM. Genetic and environmental risk factors for rheumatoid arthritis. *Best Pract Res Clin Rheumatol* 2017; **31**(1): 3-18.
23. Qvarfordt M, Andersson ML, Larsson I. Factors influencing physical activity in patients with early rheumatoid arthritis: A mixed-methods study. *SAGE Open Med* 2019; **7**: 2050312119874995.
24. Lee SW, Lee J, Moon SY, et al. Physical activity and the risk of SARS-CoV-2 infection, severe COVID-19 illness and COVID-19 related mortality in South Korea: a nationwide cohort study. *Br J Sports Med* 2022; **56**(16): 901-12.
25. Almoallim H, Al Saleh J, Badsha H, et al. A Review of the Prevalence and Unmet Needs in the Management of Rheumatoid Arthritis in Africa and the Middle East. *Rheumatol Ther* 2021; **8**(1): 1-16.
26. Cui A, Li H, Wang D, Zhong J, Chen Y, Lu H. Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. *EClinicalMedicine* 2020; **29-30**: 100587.
27. Joseph RP, Pituch KA, Guest MA, et al. Physical Activity Among Predominantly White Middle-Aged and Older US Adults During the SARS-CoV-2 Pandemic: Results From a National Longitudinal Survey. *Front Public Health* 2021; **9**: 652197.
28. Akkuzu G, Bes C, Özgür DS, et al. Inflammatory rheumatic diseases developed after COVID-19 vaccination: presentation of a case series and review of the literature. *Eur Rev Med Pharmacol Sci* 2023; **27**(5): 2143-51.
29. Aliyu AA. Public health ethics and the COVID-19 pandemic. *Ann Afr Med* 2021; **20**(3): 157-63.
30. Stommel M, Schoenborn CA. Accuracy and usefulness of BMI measures based on self-reported weight and height: findings from the NHANES & NHIS 2001-2006. *BMC Public Health* 2009; **9**: 421.

# Machine learning-based prediction of suicidality in Korean adolescents with allergic rhinitis: derivation and validation in two independent nationwide cohorts

Ho Jae Lee,[1] Hwi Yang,[1] Dong Keon Yon,[1,2]
[1]*Center for Digital Health, Medical Science Research Institute, Kyung Hee University Medical Center, Kyung Hee University College of Medicine, Seoul, South Korea,* [2]*Department of Pediatrics, Kyung Hee University Medical Center, Kyung Hee University College of Medicine, Seoul, South Korea*
**Correspondence: Dong Keon Yon (yonkkang@gmail.com)*

*Abstract*— **Our goal is to create and verify machine learning models that leverage data from two nationwide cohort studies to forecast the risk of suicide attempts. By using a variety of data fields and conducting external validations, we aim to enhance previous models, potentially resulting in more accurate risk assessment tools. We utilized secondary data from the Korea Youth Risk Behavior Web-based Survey (KYRBS) and the Korea National Health and Nutrition Examination Survey (KNHANES) to construct machine learning algorithms capable of predicting suicide attempts in adolescents with allergic rhinitis. Four well-known machine learning algorithms were applied, and the created models underwent external validation through two studies. We discovered that integrating various categories of survey data led to the best outcomes. Adaptive Boosting (ADB) demonstrated superior performance in the development cohort, with an AUROC of 84.15 (95% CI 83.86-84.44), and also excelled in the external validation cohort, with an AUROC of 77.51. By using data from two cohorts, we were able to develop and authenticate machine learning models that can predict suicide attempts in adolescents with allergic rhinitis. These models showed satisfactory performance and are potentially applicable in real-world clinical scenarios.**

## I. INTRODUCTION

Allergic Rhinitis (AR), affecting approximately 14% of the global population[1], impacts quality of life and is associated with depressive symptoms, including suicidal ideation[2,3]. With adolescent suicide rates rising, particularly in Korea,[4] and adolescents' inherent vulnerability to mental health problems, understanding AR's role in this critical issue is essential.

Suicide prediction presents significant challenges, with no established practical methods for individual suicide anticipation or risk stratification.[5,6] However, machine learning (ML) could potentially improve accuracy in identifying adolescents at risk of suicide.[7] Despite its limited application, ML has shown significant advancements in suicide prediction, indicating its promising role in precision medicine, particularly in handling large and complex databases.[8,9]

Considering the additional suicide-related risks in adolescents with AR, this study aims to utilize ML to better understand this risk. The objective is to develop and investigate a ML model predicting suicide attempts in AR patients, using two independent nationwide cohorts in South Korea, based on nationwide population data.

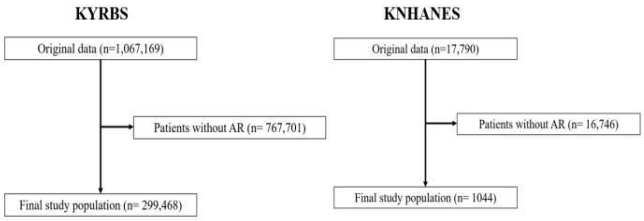



Fig.1 Study workflow

## II. METHODS

### A. Database

This study aimed to develop a machine learning model to predict suicidality in Korean adolescents using data from the Korea Youth Risk Behavior Web-based Survey (KYRBS) and the Korea National Health and Nutrition Examination Survey (KNHANES).[10] The model was trained, validated, and tested on various clinical features, including demographic, socioeconomic, and mental health factors. The goal was to improve the identification and intervention of mental health conditions among this population. The study was approved by the Institutional Review Board of the Korean Centers for Disease Control and Prevention Agency (KCDA). The study population of each study were 2099,468, and 1044 each. (Figure 1)

### B. Validation and algorithm selection

The dataset was randomly split into a base training set (80%, 492) and a base test set (20%, 124) ensuring equal distribution of difficult class patient data. Significant variables for DL occurrence were identified through logistic regression of the base training set, leading to the creation of several training sets with varying combinations of variables. Corresponding test sets were created for each of these. The

TABLE I

HYPERPARAMETERS AND THEIR VALUES IN EACH MODEL

| | Booster | Learning rate | Max depth | N estimators | Subsample | Max features | Estimator | Colsample bytree | Num leaves |
|---|---|---|---|---|---|---|---|---|---|
| ADB | N/A | 0.5 | N/A | 50 | N/A | N/A | Decision tree classifier (max depth=3) | N/A | N/A |
| RF | N/A | N/A | 10 | 100 | N/A | sqrt | N/A | N/A | N/A |
| LGB | N/A | 0.2 | 8 | 150 | 0.6 | N/A | N/A | 1 | 40 |
| XGB | GBtree | 0.1 | 20 | 200 | 0.5 | N/A | N/A | 0.5 | N/A |

objective of the study was to develop a predictive model with a minimal number of variables yet high performance. Thus, each model trained with the basic training set was compared with models having fewer variables. Each training set was normalized by min-max scaling after applying SMOTE, while each test set was normalized using the min-max scaling of the training set. Data variables were analyzed using SAS software, version 9.3.

c. *Machine Learning Model*

This study analyzed the original Adolescent Health Survey dataset, dividing it into training and test datasets (4:1 ratio). Various machine learning algorithms were applied to the training dataset and their performance was assessed using Area Under the Receiver Operating Characteristic (AUROC) scores on the test dataset. High-performing models were further examined. An additional dataset from the Korea National Health and Nutrition Examination Survey (KNHANES) was used for external validation, comparing performance metrics with the training and test datasets.

Data pre-processing involved handling missing values, encoding categorical variables, scaling numerical features, and other data transformations to prepare the data for analysis. A 10-fold cross-validation approach was used to reliably assess model performance. The 95% Confidence Intervals (CI) were calculated for each performance metric during the 10-fold cross-validation to estimate the uncertainty and variability of the results.

Data processing was performed using SAS, while machine learning analysis utilized Python, TensorFlow-gpu, Keras, NumPy, Pandas, Scikit-learn, and Matplotlib. The machine learning models used were decision tree model AdaBoost, Random Forest, Light Gradient Boost, and Xgboost. Hyperparameters were fine-tuned using GridSearch to maximize the AUROC scores. Feature importance was also determined to understand the influence of features on the prediction of the target variable.

## III. RESULTS

### *A. Demographic Characteristics*

The study was conducted using nationwide population data from two independent cohorts in South Korea to develop and investigate a machine-based learning model for predicting suicide attempts in patients with allergic rhinitis. The study population's demographic characteristics were as follows

Both cohorts consisted of a significant number of patients with allergic rhinitis in the age range of 13 to 18 years, which corresponds to the adolescent age group. In Cohort 1, the gender distribution revealed that 51.02% were males and 48.98% were females. In Cohort 2, 59.87% were males and 40.13% were females. The patient samples in both cohorts encompassed diverse socioeconomic backgrounds, including varying levels of education, income, and occupations.

Overall, the study included a substantial number of patients with allergic rhinitis from diverse demographic backgrounds, ensuring a representative sample for the development and evaluation of the machine learning model. By considering these demographic characteristics, the study aimed to provide valuable insights into the risk of suicide attempts among individuals with allergic rhinitis, with a particular focus on the adolescent population.

### *B. Machine Learning Model Results*

The best parameters of machine learning algorithms in predicting in predicting suicide attempt with AR is presented in Table 1. The performance of the machine learning algorithms is presented in Table 2. Adaboost was the best

TABLE 2

STUDY DATA VALIDATION, TEST RESULTS, AND EXTERNAL VALIDATION DATA RESULTS

| | Model | Sensitivity, % | Specificity, % | Accuracy, % | Balanced accuracy, % | AUROC, % |
|---|---|---|---|---|---|---|
| Validation results (95% CI) | **ADB** | **76.58(76.28-76.68)** | **76.48(76.29-76.67)** | **76.48(76.29-76.67)** | **76.48(76.28-76.67)** | **84.15(83.86-84.44)** |
| | RF | 76.19(75.86-76.52) | 76.18(75.87-76.49) | 76.18(75.87-76.49) | 76.18(75.86-76.51) | 84.09(83.86-84.33) |
| | LGB | 75.80(75.42, 76.18) | 75.81(75.43, 76.18) | 75.81(75.43, 76.18) | 75.80(75.43, 76.18) | 83.38(83.13, 83.64) |
| | XGB | 74.17(73.69, 74.64) | 74.18(73.69, 74.67) | 74.18(73.69, 74.67) | 74.17(73.69, 74.66) | 81.32(80.80, 81.83) |
| Test results | **ADB** | **76.31** | **76.30** | **76.30** | **76.30** | **83.84** |
| | RF | 72.62 | 75.80 | 75.80 | 75.81 | 83.76 |
| | LGB | 75.34 | 76.14 | 76.11 | 75.74 | 83.28 |
| | XGB | 73.98 | 73.96 | 73.96 | 73.97 | 81.16 |
| External results1 | **ADB** | **66.67** | **74.03** | **73.95** | **70.35** | **77.51** |

model in predicting suicide attempts in allergic rhinitis patients. AdaBoost resulted in validation results with a sensitivity of 76.58% (95% CI: 76.28-76.68), specificity of 76.48% (95% CI: 76.29-76.67), accuracy at 76.48% (95% CI: 76.29-76.67), balanced accuracy of 76.48% (95% CI: 76.28-76.67), and AUROC of 84.15% (95% CI: 83.86-84.44). (Figure 1) Random forest had validation results with a sensitivity of 76.19% (95% CI: 75.86-76.52), specificity of 76.18% (95% CI: 75.87-76.49), accuracy at 76.18% (95% CI: 75.87-76.49), balanced accuracy of 76.18% (95% CI: 75.86-76.51), and AUROC of 84.09% (95% CI: 83.86-84.33) (Figure 1) Light Gradient Boost showed a sensitivity of 75.80% (95% CI: 75.42-76.18), specificity of 75.81% (95% CI: 75.43-76.18), accuracy at 75.81% (95% CI: 75.43-76.18), balanced accuracy of 75.80% (95% CI: 75.43-76.18), and finally AUROC of 83.38% (95% CI:83.13-83.64). (Figure 1) XGB had validation results with a sensitivity of 74.17% (95% CI: 73.69-74.64), specificity of 74.18% (95% CI: 73.69-74.67), accuracy at 74.18% (95% CI:73.69-74.67), balanced accuracy of 74.17% (95% CI: 73.69-74.67), and AUROC of 81.32% (95% CI: 80.80-81.83).

### *C. Extra validation results*

We externally validated the model using adolescent data from KNHANES to evaluate the model performance with different dataset. The ADB model predicted suicidality with an AUROC of 77.51%, a sensitivity of 66.67%, and a specificity of 74.03%; Table 1).

## IV. CONCLUSION

Our study highlights the utility of machine learning models, specifically the Adaboost model, in predicting suicide attempts among Korean adolescents with allergic rhinitis. Key predictive factors included age, academic achievement, BMI, stress, and economic status, emphasizing the importance of considering atopic conditions, particularly atopic dermatitis, in suicidality risk assessment. Despite the study's focus on a specific population, these findings suggest a broader application of machine learning in suicide risk evaluation, advocating for more comprehensive screening among at-risk adolescents, and stressing the need for public awareness of untreated allergic diseases' impact on mental health.

## V. ACKNOWLEDGEMENT

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI22C1976).

## REFERENCES

1. D'Amato, G., *et al.* Meteorological conditions, climate change, new emerging factors, and asthma and related allergic disorders. A statement of the World Allergy Organization. *The World Allergy Organization journal* **8**, 25 (2015).

2. Tenero, L.*, et al.* Diagnosis and Management of Allergic Rhinitis in Asthmatic Children. *Journal of asthma and allergy* **16**, 45-57 (2023).
3. Bell, I.R., Jasnoski, M.L., Kagan, J. & King, D.S. Depression and allergies: survey of a nonclinical population. *Psychotherapy and psychosomatics* **55**, 24-31 (1991).
4. Lee, S.U.*, et al.* Changing trends in suicide rates in South Korea from 1993 to 2016: a descriptive study. *BMJ open* **8**, e023144 (2018).
5. Franklin, J.C.*, et al.* Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological bulletin* **143**, 187-232 (2017).
6. Large, M.*, et al.* Meta-Analysis of Longitudinal Cohort Studies of Suicide Risk Assessment among Psychiatric Patients: Heterogeneity in Results and Lack of Improvement over Time. *PloS one* **11**, e0156322 (2016).
7. McHugh, C.M. & Large, M.M. Can machine-learning methods really help predict suicide? *Current opinion in psychiatry* **33**, 369-374 (2020).
8. Bernert, R.A.*, et al.* Artificial Intelligence and Suicide Prevention: A Systematic Review of Machine Learning Investigations. *Int J Environ Res Public Health* **17**(2020).
9. Burke, T.A., Ammerman, B.A. & Jacobucci, R. The use of machine learning in the study of suicidal and non-suicidal self-injurious thoughts and behaviors: A systematic review. *Journal of affective disorders* **245**, 869-884 (2019).
10. Kim, W., Chun, S. & Lee, S.A. Suicide attempt and violence victimization in Korean adolescents with migrant parents: A nationwide study. *Journal of affective disorders* **290**, 164-168 (2021).

# Design of Evolutionary Algorithm based Unequal Clustering for Energy Aware Wireless Sensor Networks

Mohammed Altaf Ahmed[1], T. Satyanarayana Murthy[2], Fayadh Alenezi[3], E. Laxmi Lydia[4], Seifedine Kadry[5], Yena Kim[6] and Yunyoung Nam[6],

*[1]*Department of Computer Engineering, College of Computer Engineering & Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia*

[2]*Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India.*

[3]*Department of Electrical Engineering, College of Engineering, Jouf University, Saudi Arabia*

[4]*Department of Computer Science and Engineering, Vignan's Institute of Information Technology, Visakhapatnam, 530049, India*

[5]*Department of Applied Data Science, Noroff University College, Kristiansand, Norway*

[6]*Department of ICT Convergence, Soonchunhyang University, South Korea*

[*]*Corresponding Author: Yunyoung Nam. Email: ynam@sch.ac.kr*

*Abstract*— **Wireless Sensor Networks (WSN) plays a vital role in several real time applications ranging from military to civilian applications. Despite the benefits of WSN, energy efficiency becomes a major part of challenging issue in WSN, which necessitate proper load balancing amongst the clusters and serves a wider monitoring region. Clustering technique for WSN has several benefits namely lower delay, higher energy efficiency, and collision avoidance. But clustering protocol has number of challenges. In a largescale of network, cluster-based protocols mainly adapt multi-hop routing for saving energy that leads to hot spot problems. A hot spot problem becomes a problem where cluster node nearer to the base station (BS) tends to drain the energy much quicker than other nodes because of the need to implement more transmission. This article introduces a Jumping Spider Optimization Based Unequal Clustering Protocol for Mitigating Hotspot Problem (JSOUCP-MHP) in WSN. The JSO algorithm is stimulated by the characteristics of spiders naturally and mathematically modelled the hunting mechanism such as search, persecution, and jumping skills to attack prey. The presented JSOUCP-MHP technique mainly resolves the hot spot issue for maximizing the network lifespan. To attain this, the JSOUCP-MHP technique elects proper set of cluster heads (CHs) using average residual energy (RE). In addition, the JSOUCP-MHP technique determines the cluster sizes based on two measures i.e., RE and distance to BS (DBS). The proposed JSOUCP-MHP technique is examined under several experiments to ensure its supremacy. The comparison study shows the significance of the JSOUCP-MHP technique over other models.**

## I. INTRODUCTION

Wireless Sensor Networks (WSNs) are advanced for monitoring and sensing vital signs of area and environment by making use of linked and distributed sensor nodes (SNs). The SNs were divided into Gateway Nodes (GN), normal, and sink nodes [1]. The SNs were small in size and have insufficient sources with regard to processing, storage, and space energy [2]. In many cases, the SNs were deployed in a very harsh and intense ecosystem. The first and foremost goal of this positioning was sense information distantly and sends it to system or user for decision making. In order to send data, there comes a demand for more effective systems for managing the energy system of nodes and enhancing network lifetime [3]. The SNs perceiving and sensing the data from nearby environment and making processes and transmit it to the nearby nodes until data arrived at the base station (BS). In WSN, because of the restricted energy sources of SNs, there comes a crucial need for well-effective and balanced data aggregation system and energy efficient routing protocol [4]. The energy feature becomes one such factor to be considered for devising any solution for WSNs. Numerous types of routing protocols were devised for conserving the energy of SNs [5]. The clustering approach becomes an effective topological control system that could efficiently enhance the scalability duration and lifetime of WSNs. The WSN applications have achieved popularity including health monitoring, target tracking, disaster response, environmental monitoring, and security [6].

In a clustering ecosystem, data can be involved from one node to another, and occurs energy holes or hotspot difficulties [7]. A hotspot can be made by SNs positioned nearby the BS and rapidly drain energy because traffic arises from other nodes and is sent from it. Such SNs not just transmit their data but also sends the data from other bases because the initial death of nodes causes hotspot problems. The methods namely unequal clustering methods and mobile sinks or mobile data mules were majorly utilized for resolving this type of problem [8]. But many unequal clustering techniques devised for the solution of the hotspot leverages cluster which smaller in size adjacent to BS and cluster size rises case when we were going far away from BS. The cluster size was inversely proportional to the distance from BS. The clusters adjacent to BS hold a greater quantity of nodes that is helpful in efficient load sharing [9]. The projected method for size allotment and size difference of clusters leads to reduction of frequency where a specific node turns out to be cluster head (CH) [10]. This aids in maintenance of the overall connectivity and thwarts network isolation. In line with this, hot spot issues can be reduced. One of the problem with regard to unequal clustering was cluster size decreasing or

increasing ratio that was not evidently conferred in prevailing methods.

Arikumar et al. [11] modelled an Energy Efficient LifeTime Maximization (EELTM) technique that uses intelligent techniques Fuzzy Inference System (FIS) and particle swarm optimization (PSO). In addition, devised an optimal CH–CR selecting method in this technique that exploits fitness values (FV) computed by the PSO approach for determining 2 optimally nodes in every cluster to serve as Cluster Router (CR) and CH. The CH which is selected exclusively accumulates the data from its cluster members, although the CR is accountable to receive the data which is collected from its CH and sending it to BS. Mehra [12] formulated an enhanced fuzzy unequal clustering and routing protocol (E-FUCA) in which vital variables were regarded at the time of CH candidate selecting, and intellectual decision utilizing FL was considered by non-CH nodes at the time of election of its CH for forming clusters. To expand the lifespan, FL is utilized as next-hop choice for effective routing. And carried out the simulation experimentations for 4 cases and comparison is made with the propound performance of protocol with recent similar protocols.

Wang and Hu [13] introduced an energy-efficient unequal clustering routing algorithm (UCRA). At first, the monitoring region has been separated by concentric circles as rings of various sizes. Then, the CH was selected relating to RE and position. Before clustering, every typical sensor links cluster related to the electability of every CH. Nguyen et al. [14] presented a new optimization technique, like the compact bat algorithm (cBA), for using it for class of optimization issues which involves gadgets containing inadequate hardware sources. A real-valued prototype vector can be employed for probabilistic functions for generating every candidate solution for optimizing cBA. The projected cBA was widely assessed on numerous continual multi-modal functions and the unequal clustering of WSN (uWSN) issues.

In [15], competitive swarm optimization (CSO) related methods were introduced, together called such methods as CSO-UCRA. Initially, the CH selecting approach was provided relies on CSO related approach, after assigning non-CH sensors to CHs. At last, a CSO related routing method was provided. New fitness functions and efficient particle encoding techniques were advanced for such methods. Guleria and Verma [16] devise the new ant colony meta-heuristic related unequal clustering for novel CH election. The neighbor finding level and link maintenance by Meta-Heuristic Ant Colony Optimization (ACO) technique choose the optimal path among the nodes that rises the packets distributed to the destiny.

This article introduces a Jumping Spider Optimization Based Unequal Clustering Protocol for Mitigating Hotspot Problem (JSOUCP-MHP) in WSN. The JSO algorithm is stimulated by the characteristics of spiders in nature and mathematically modelled the hunting mechanism such as search, jumping skills, and persecution to attack prey. The presented JSOUCP-MHP technique mainly resolves the hot spot issue for maximizing the network lifespan. To attain this, the JSOUCP-MHP technique elects proper set of cluster heads (CHs) using average RE. In addition, the JSOUCP-MHP technique determines the cluster sizes based on two measures i.e., RE and distance to BS (DBS). The proposed JSOUCP-MHP technique is examined under several experiments to ensure its supremacy.

## II. The Proposed Unequal Clustering Protocol

In this article, a novel JSOUCP-MHP algorithm was presented for resolving hot spot issues in WSN. The JSO algorithm is stimulated by the characteristics of spiders in nature and mathematically modelled the hunting mechanism such as search, jumping skills, and persecution to attack prey. The presented JSOUCP-MHP technique mainly resolves the hot spot issue for maximizing network lifespan. Fig. 1 depicts the overall process of JSOUCP-MHP approach.

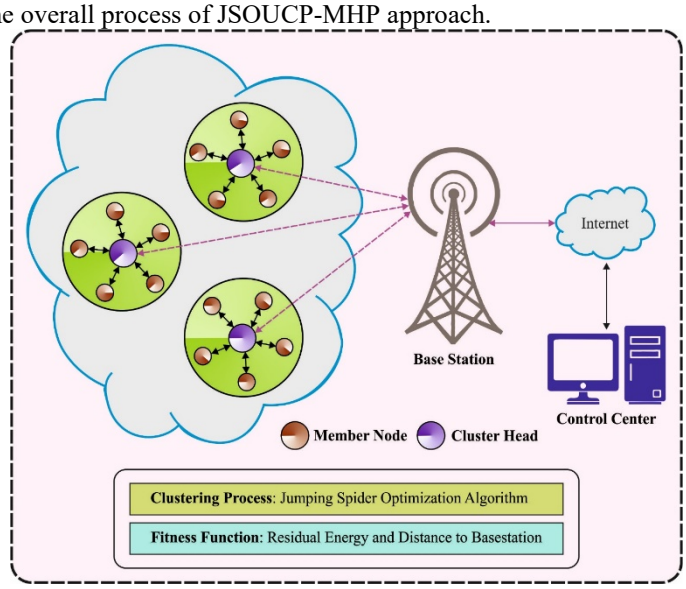



Fig. 1 Overall process of JSOUCP-MHP approach

### A. System Model

A WSN comprises N uniform distribution sensors. There is one BS interconnected with users through the Internet. Let assume $S = \{S_1, S_{2,\dots}, s_i, s_{N-1}, s_N\}$ as the group of nodes, whereby s_i denotes a nodei and |S|=N. WNS makes use of wireless radio transceiver depending on the different variables, for example, energy utilization and distance. The distance among the receiver and transmitter followed by attenuated transceiver power exponentially reduced with increased distance. Now, we represent (x_i; y_i) and (x_j; y_j) were the coordinates of nodes i and j. Assume node i sent to destination node j, the energy consumed on communication over d distance can be evaluated as follows:

$$E_{\mathrm{T}x_j(l,d)} = E_{T_{X-elec}}(l) + E_{\mathrm{T}x-amp}(l,d) = \begin{cases} I \times Eelec + \varepsilon_{fs} \times d^2, d < d_0 \\ I \times Eelec + \varepsilon_{mp} \times d^4, d \geq d_0 \end{cases} \quad (1)$$

In Eq. (1), E refers to the energy consumption; Tx is to transfer, elect, amp specifies electronic and amplify to modulate, digitalized coding, spreading, and filter signals. It illustrates that the multi - hop transmission is highly efficient in WSN. $d_o$ denotes a threshold of space model,

$$d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}}, \quad (2)$$

In Eq. (2), the power loss ε_(fs/) ε_mp are free space and multi-path model and it is formulated by:

$$E_{Rxij(l)} = E_{R_{X-elec}}(l) = l \times E_{elec}. \quad (3)$$

WSN considered N node in the region of 2D M^2 with k cluster. In relation to the hotspot perplexity in WSN was balancing the load amongst CHs according to the clustering

technique. Unequal clustering decreases the cluster size nearer to BS and the size of cluster rises as distance among the BS ad well as CH rises. The cluster member senses the real-time parameter and transfers the sensed value to CH. CH aggregates and receives information to eliminate inessential information and transfer aggregated information to BS directly or through intermediatory CH. In equal clustering, the cluster size remains unchanged all over the network. But, in unequal clustering, the cluster size can be described according to the DBS [17]:

$$E_{c1uster} = E_{CH} + \left(\frac{N}{k} - 1\right) \times E_{member,} \quad (4)$$

Let $N/k - 1$ be the average of member nodes from cluster, $E_{CH}$, $E_{members}$ denotes dissipated energy for CH and member correspondingly and it is evaluated by:

$$E_{member} = l \times E_{e1ec} + l \times \varepsilon_{fs} \times d_{toCH}^2, \quad (5)$$

$$E_{CH} = l \times E_{e1ec} \times \left(\frac{N}{k} - 1\right) + l \times E_{DA} \times \frac{N}{k} + l \times E_{e1ec} + l \times \varepsilon_{mp} \times d_{toBS}^4. \quad (6)$$

As a result, the energy consumed for WSN from the generation formulated from energy process and transceiver was given by:

$$E_{total} = E_p + E_{frame} = E_p + k \times E_{c1uster}, \quad (7)$$

In Eq. (7), $E_p$ is energy consumed for microcontroller and source voltage of node. It doesn't affect the optimization process. As a result, the consumed energy $E_{frame}$ optimizes according to the distance for clustering optimization.

### *B. Design of JSO Algorithm*

The mathematical modelling of jumping spider's hunting strategy is initially proposed. Then, The JSOA approach was introduced. The hunting strategies are searched, jumping on the prey, and attacking by persecution [18]. Also, the algorithm describes the pheromone rate of the spider as follows.

Once the spider is away from a distance whereas it could catch the prey by jumping, it moves closer by indulging certain stealthy movement till it was at an attainable distance where it could catch the prey and pounces on it. The persecution approach can be described as the uniform accelerated rectilinear motion.

$$x_i = \frac{1}{2}at^2 + v_0 t \quad (8)$$

In Eq. (8), $x_i$ demonstrates the location of $i^{th}$ follower spider, $t$ refers to the time, $v_0$ represent the speed initially. The acceleration can be shown as $a = \frac{v}{t}$, where $v = \chi - \chi_0$.

In this study, to optimize, every iteration can be regarded as the duration, in which the differences between iterations are equivalent to 1 and initially the speed is fixed as zero, $v_0 = 0$. And it is formulated as follows:

$$\chi_{i(g+1)=} \frac{1}{2}\left(\overrightarrow{x_i}(g) - x_r \rightarrow (g)\right) \quad (9)$$

In Eq. (9), $\overrightarrow{x_i}(g+1)$ represents the novel location of search agent to generation $g+1$, $\overrightarrow{x_i}(g)$ indicates the present $i\text{-}th$ searching agent in generation $g$, and $\overrightarrow{x_r}(g) \rightarrow$ denotes the $r\text{-}th$ searching agent arbitrarily chosen, with $i \neq r$, whereas $r$ represents an arbitrary value ranging from 1 to the size of maximal searching agent.

The jumping spider follows the prey and jumps on them. The hunting strategy of jumping on prey is characterized by a projectile motion.

$$\overrightarrow{\chi_i} = v_0 \cos(\alpha)\overrightarrow{ti}$$
$$\frac{dx}{dt} = \overrightarrow{V_x} = v_0 \cos(\alpha)\overrightarrow{\iota} \quad (10)$$

Likewise, vertical axis and the derivative w.r.t time can be given as.

$$\overrightarrow{y_i} = (v_0 sen(\alpha)t_2^1 - gt^2)\vec{j} \quad (11)$$
$$\frac{dy}{dt} = \overrightarrow{V_y} = (v_0 sen(\alpha) - gt)\vec{j}$$

Similarly, time is characterized by strategy 1. Thus, we attain the trajectory equation as follows.

$$y = x\tan(\alpha) - \frac{\overrightarrow{gx_i^2}}{2V_0^2\cos^2(\alpha)} \quad (12)$$

Lastly, the trajectory is formulated by:

$$\overrightarrow{x_i}(g+1) = \overrightarrow{x_i}(g)\tan(\alpha) - \frac{\overrightarrow{gx^2}(g)}{2V_0^2\cos^2(\alpha)} \quad (13)$$
$$\alpha = \frac{\varphi\pi}{180}$$

The Jumping spider implements a random search around the environment for locating prey. Local and global search are the two mathematical functions presented as follows.

The local search can be defined as follows:

$$\overrightarrow{x_i}(g+1) = x_{best}(g) + walk \rightarrow (\tfrac{1}{2} - \varepsilon) \quad (14)$$

In Eq. (14), $\overrightarrow{x_i}(g+1)$ refers to novel location of searching agent, $\overrightarrow{x_{best}}(g)$ denotes the optimal searching agent found from the preceding iteration, walk was a normally distributed pseudo-random integer within $(-2,2)$, $\varepsilon$ denotes a pseudo-random number uniformly distributing from $(0,1)$.

At the same time, the Global search can be expressed as follows.

$$\overrightarrow{x_i}(g+1) = \overrightarrow{x_{best}}(g) + \left(\overrightarrow{x_{best}}(g) - \overrightarrow{x_{worst}}(g)\right)\lambda \quad (15)$$

In Eq. (15), $\overrightarrow{x_i}(g+1)$ denotes the novel location of a searching agent, $\overrightarrow{x_{best}}(g)$ and $\overrightarrow{x_{worst}}(g)$ indicates the best and worst searching agent form preceding iteration, correspondingly, and $\lambda$ refers to a Cauchy random integer with $\mu$ fixed as 0 and $\theta$ fixed as 1.

Pheromone is released by several animals, amongst which were insects, together with spiders. Nonetheless, they produce pheromone; the modelling of rate of pheromone was taken and described as follows:

$$pheromone\ (i) = \frac{Fitness_{\max} - Fitness(i)}{Fitness_{\max} - Fitness_{\min}} \quad (16)$$

In Eq. (16), $Fitness_{\max}$ and $Fitness_{min}$ represent worst and the best FV in present generation, correspondingly, while Fitness(i) refers to the present FV of $i\text{-}th$ searching agent. Eq. (16) normalizes the FV within zero and one whereas 0 indicates the worst pheromone rate, while 1 denotes the optimal. The criteria involve for lower pheromone rate value equivalent or lesser than 0.3 as follows:

$$\overrightarrow{x_i}(g) = \overrightarrow{x_{best}}(g) + \frac{1}{2}(\overrightarrow{x_{r_1}}(g) - (-1)^{\sigma\overrightarrow{x_{r_2}}}(g)) \quad (17)$$

In Eq. (17), $\overrightarrow{x_i}(g)$ represent the searching agent (jumping spider) with lower pheromone rate that is upgraded, $r_1$ and $r_2$ indicates arbitrary number produced from 1 to the maximal size of searching agent, with $r_1 \neq r_2$, while $\overrightarrow{x_{r_1}}(g)$ and $\overrightarrow{x_{r_2}}(g)$ are

the $r_1$, $r_2th$ searching agent selected, $\overrightarrow{x_{best}}(g)$ indicates the best searching agent from preceding iteration and $\sigma$ indicates a binary number produced, $\sigma \in \{0,1\}$.

### *C. Process Involved in JSOUCP-MHP Technique*

In real-time application of WSN, a massive amount of nodes is positioned generally with a higher node density. Assume a proper threshold $T$ for controlling the proportion of $CH$, $T$ is fixed to 0.4. Every node $S_j$ evaluates the value of $\mu$. When $\mu < T$, node $S_j$ become a CCH; or else, node $S_j$ become a regular node and go to a dormant state until the last CH selection is accomplished as follows.

$$\mu = \mu_0 \frac{E_{avg}}{RE_i}, \quad (18)$$

In Eq. (18), $\mu_0$ refers to a uniformly distributed arbitrary value within [0,1] and $E_{avg}$ indicates the average RE of alive node. The computation method was given in the following: The data packet transferred by the node includes $RE_i$. Afterward, the BS received, the average RE $E_{avg}$ of each alive node is evaluated; $RE_i$ characterizes the RE of node $S_i$. Clearly, the large of $RE_i$, the small of $\mu$, hence the better possibility that $S_i$ becomes a CH.

In JSOUCP-MHP technique, consider the distance from CH to BS and ignore the energy. This might leads to lower energy node to be the CH, and afterward generating a cluster, the member in the cluster causes an increase in the energy usage of CH. As a result, a competitive radius considering distance and energy is presented, and the competitive radius for CH $v_i$ is evaluated as follows.

$$v_i.R_{comp} = \left[1 - w_1 \frac{d_{\max} - d(v_i,BS)}{d_{\max} - d_{\min}} - w_2 \left(1 - \frac{RE_{v_i}}{E_0}\right)\right] \times R^0_{comp}, \quad (19)$$

In Eq. (5), $d_{\max}$ and $d_{\min}$ denotes the maximal and minimal distances from the alive node to the $BS$, correspondingly, $d(v_i, BS)$ represents the distance from $v_i$ to the $BS$, $RE_{v_i}$ indicates the $RE$ of $v_i$, $w_1$ and $w_2$ are constant within [0,1], and $w_1 + w_2 = 1$.

## III. PERFORMANCE VALIDATION

A brief set of experimental analyses is carried out to demonstrate the enhanced performance of the JSOUCP-MHP model on WSN.

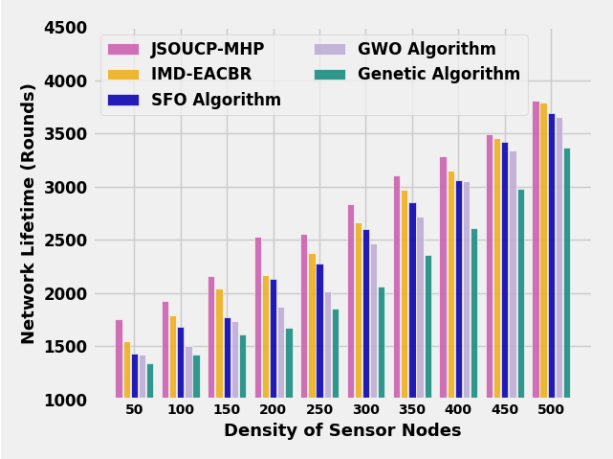
Tab. 1 and Fig. 2 showcase the network lifetime (NLT) inspection of the JSOUCP-MHP model with existing models under varying density of sensor nodes (DSN). The experimental outcomes depicted that the JSOUCP-MHP model has shown enhanced performance with higher NLT values. For instance, with 50DSN, the JSOUCP-MHP model has exhibited increased NLT of 1757 whereas the enhanced metaheuristic-driven energy-aware cluster-based routing (IMD-EACBR), sunflower optimization (SFO), grey wolf optimization (GWO), and genetic algorithm (GA) models have demonstrated reduced NLT of 1549, 1433, 1422, and 1343 respectively. Moreover, with 250DSN, the JSOUCP-MHP model has attained higher NLT of 2556 whereas the IMD-EACBR, SFO, GWO, and GA models have obtained lower NLT of 2374, 2276, 2012, and 1852 respectively. On the other hand, with 500DSN, the JSOUCP-MHP model has illustrated improved NLT of 3802 whereas the IMD-EACBR, SFO, GWO, and GA models have depicted reduced NLT of 3789, 3691, 3650, and 3369 respectively.

TABLE I

NLT ANALYSIS OF JSOUCP-MHP APPROACH WITH RECENT ALGORITHMS UNDER DISTINCT DSNS

Fig. 2 NLT analysis of JSOUCP-MHP approach under distinct DSNs

Tab. 2 and Fig. 3 exhibits the number of Alive Sensor Nodes (NOASN) analysis of the JSOUCP-MHP algorithm with existing models under varying count of rounds. The experimental outcomes depicted that the JSOUCP-MHP approach has shown enhanced performance with higher NOASN values. For example, with 2000 rounds, the JSOUCP-MHP approach has exhibited improved NOASN of 481 where the IMD-EACBR, SFO, GWO, and GA methods have




demonstrated reduced NOASN of 473, 405, 391, and 283 correspondingly. Furthermore, with 3000 rounds, the JSOUCP-MHP method has reached higher NLT of 367 whereas the IMD-EACBR, SFO, GWO, and GA models have acquired lower NOASN of 264, 112, 80, and 78 correspondingly. In contrast, with 4000 rounds, the JSOUCP-MHP model has demonstrated improved NOASN of 176 whereas the IMD-EACBR, SFO, GWO, and GA models have shown reduced NOASN of 87, 11, 0, and 0 correspondingly.

TABLE II

NOASN ANALYSIS OF JSOUCP-MHP APPROACH WITH RECENT ALGORITHMS UNDER DISTINCT ROUNDS

| No. of Alive Sensor Nodes (NOASN) | | | | | |
|---|---|---|---|---|---|
| No. of Rounds | JSOUCP-MHP | IMD-EACBR | SFO Algorithm | GWO Algorithm | Genetic Algorithm |
| 2000 | 481 | 473 | 405 | 391 | 283 |
| 2250 | 468 | 386 | 343 | 267 | 182 |
| 2500 | 451 | 393 | 230 | 216 | 125 |
| 2750 | 435 | 317 | 175 | 135 | 96 |
| 3000 | 367 | 264 | 112 | 80 | 78 |
| 3250 | 289 | 168 | 86 | 59 | 41 |
| 3500 | 245 | 133 | 49 | 26 | 19 |
| 3750 | 204 | 109 | 19 | 0 | 0 |
| 4000 | 176 | 87 | 11 | 0 | 0 |

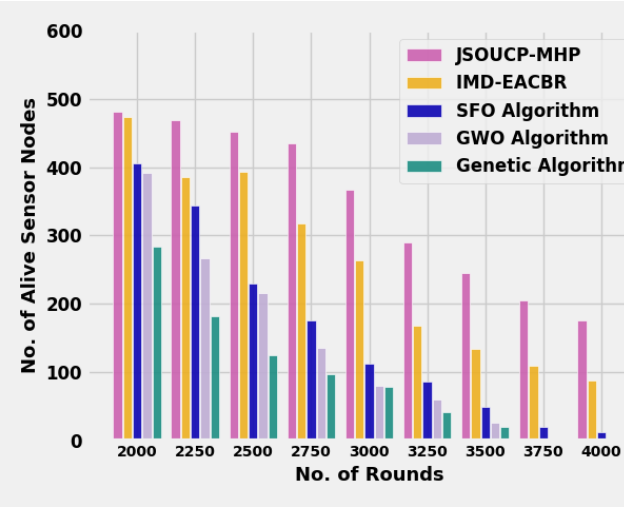



Fig. 3 NOASN analysis of JSOUCP-MHP approach under distinct rounds

A detailed number of dead sensor node (NODSN) assessments of the JSOUCP-MHP with recent approaches were performed in Tab. 3 and Fig. 4. The results inferred the JSOUCP-MHP technique has resulted in enhanced results with minimum values of NODSN. For instance, with 2000 rounds, the JSOUCP-MHP model has achieved least NODN of 19 whereas the IMD-EACBR, SFO, GWO, and GA models have reached increased NODN of 27, 95, 109, and 217 respectively. Meanwhile, with 3000 rounds, the JSOUCP-MHP model has resulted to decreased NODN of 133 whereas the IMD-EACBR, SFO, GWO, and GA models have exhibited increased NODN of 236, 388, 420, and 422 respectively. Eventually, with 4000 rounds, the JSOUCP-MHP method has rendered minimal NODN of 324 whereas the IMD-EACBR, SFO, GWO, and GA models have achieved maximum NODN of 413, 489, 500, and 500 correspondingly.

TABLE III

NODSN ANALYSIS OF JSOUCP-MHP APPROACH WITH RECENT ALGORITHMS UNDER DISTINCT ROUNDS

| No. of Dead Sensor Nodes (NODSN) | | | | | |
|---|---|---|---|---|---|
| No. of Rounds | JSOUCP-MHP | IMD-EACBR | SFO Algorithm | GWO Algorithm | Genetic Algorithm |
| 2000 | 19 | 27 | 95 | 109 | 217 |
| 2250 | 32 | 114 | 157 | 233 | 318 |
| 2500 | 49 | 107 | 270 | 284 | 375 |
| 2750 | 65 | 183 | 325 | 365 | 404 |
| 3000 | 133 | 236 | 388 | 420 | 422 |
| 3250 | 211 | 332 | 414 | 441 | 459 |
| 3500 | 255 | 367 | 451 | 474 | 481 |
| 3750 | 296 | 391 | 481 | 500 | 500 |
| 4000 | 324 | 413 | 489 | 500 | 500 |

| Network Lifetime (Rounds) | | | | | |
|---|---|---|---|---|---|
| Density of Sensor Nodes | JSOUCP-MHP | IMD-EACBR | SFO Algorithm | GWO Algorithm | Genetic Algorithm |
| 50 | 1757 | 1549 | 1433 | 1422 | 1343 |
| 100 | 1923 | 1789 | 1679 | 1504 | 1416 |
| 150 | 2160 | 2038 | 1773 | 1733 | 1611 |
| 200 | 2530 | 2171 | 2133 | 1870 | 1668 |
| 250 | 2556 | 2374 | 2276 | 2012 | 1852 |
| 300 | 2832 | 2666 | 2599 | 2468 | 2062 |
| 350 | 3104 | 2967 | 2852 | 2715 | 2354 |
| 400 | 3282 | 3145 | 3059 | 3050 | 2613 |
| 450 | 3487 | 3452 | 3418 | 3340 | 2976 |
| 500 | 3802 | 3789 | 3691 | 3650 | 3369 |

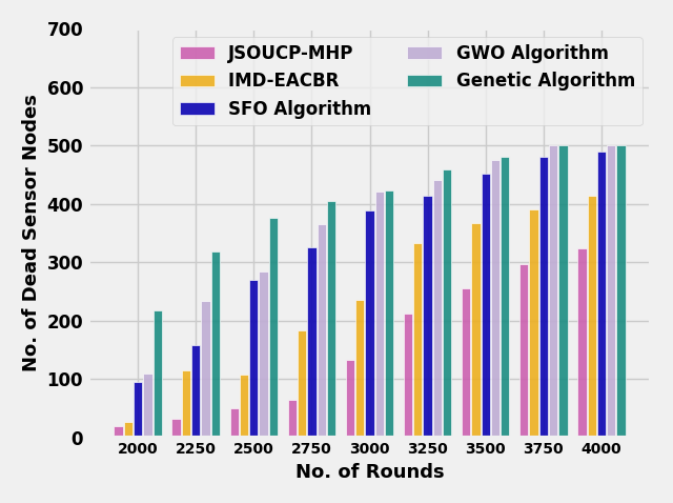



Fig. 4 NODSN analysis of JSOUCP-MHP approach under distinct rounds

A detailed energy consumption (ECON) evaluation of the JSOUCP-MHP with recent methods is executed in Tab. 4 and Fig. 5. The results implicit the JSOUCP-MHP method have resulted to enhanced results with least values of ECON. For example, with 50 DSN, the JSOUCP-MHP method del has gained least ECON of 0.0585mJ whereas the IMD-EACBR, SFO, GWO, and GA models have acquired increased ECON of 0.0936mJ, 0.1117mJ, 0.1373mJ, and 0.2183mJ correspondingly. In the meantime, with 250 DSN, the JSOUCP-MHP algorithm has resulted in decreased ECON of 0.3004mJ whereas the IMD-EACBR, SFO, GWO, and GA algorithms have exhibited improved ECON of 0.3201mJ, 0.4321mJ, 0.5201mJ, and 0.5545mJ respectively. Finally, with 500DSN, the JSOUCP-MHP model has presented minimal ECON of 0.5437mJ whereas the IMD-EACBR, SFO, GWO, and GA

models have gained maximum NODN of 0.6608mJ, 0.6678mJ, 0.8210mJ, and 0.8433mJ correspondingly.

TABLE IV

ECON ANALYSIS OF JSOUCP-MHP APPROACH WITH RECENT ALGORITHMS UNDER DISTINCT DSNS

| Energy Consumption (mJ) | | | | | |
|---|---|---|---|---|---|
| Density of Sensor Nodes | JSOUCP-MHP | IMD-EACBR | SFO Algorithm | GWO Algorithm | Genetic Algorithm |
| 50 | 0.0585 | 0.0936 | 0.1117 | 0.1373 | 0.2183 |
| 100 | 0.1199 | 0.1467 | 0.2031 | 0.2171 | 0.3396 |
| 150 | 0.1074 | 0.1851 | 0.2507 | 0.3078 | 0.4064 |
| 200 | 0.2350 | 0.2729 | 0.3578 | 0.4194 | 0.5030 |
| 250 | 0.3004 | 0.3201 | 0.4231 | 0.5201 | 0.5545 |
| 300 | 0.3910 | 0.4030 | 0.5945 | 0.5585 | 0.6187 |
| 350 | 0.4355 | 0.4941 | 0.5978 | 0.6704 | 0.6951 |
| 400 | 0.5019 | 0.5108 | 0.6196 | 0.6940 | 0.7837 |
| 450 | 0.5228 | 0.5598 | 0.6437 | 0.7575 | 0.8135 |
| 500 | 0.5437 | 0.6088 | 0.6678 | 0.8210 | 0.8433 |

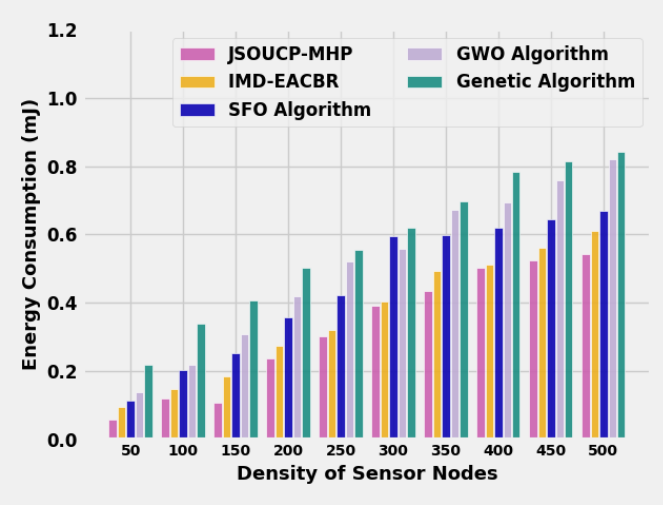



Fig. 5 ECON analysis of JSOUCP-MHP approach under distinct DSNs

Tab. 5 and Fig. 6 display the throughput (THROU) review of the JSOUCP-MHP model with existing models under varying DSN. The experimental outcomes indicated the JSOUCP-MHP model has shown enhanced performance with higher THROU values. For example, with 50DSN, the JSOUCP-MHP method has exhibited increased THROU of 0.9875Mbps whereas the IMD-EACBR, SFO, GWO, and GA models have demonstrated reduced THROU of 0.9323Mbps, 0.8915Mbps, 0.8640Mbps, and 0.8136Mbps correspondingly. In addition, with 250DSN, the JSOUCP-MHP model has gained higher THROU of 0.9517Mbps whereas the IMD-EACBR, SFO, GWO, and GA models have obtained lower THROU of 0.8955Mbps, 0.7997Mbps, 0.7435Mbps, and 0.7031Mbps correspondingly.

TABLE V

THROU ANALYSIS OF JSOUCP-MHP APPROACH WITH RECENT ALGORITHMS UNDER DISTINCT DSNS

| Throughput (Mbps) | | | | | |
|---|---|---|---|---|---|
| Density of Sensor Nodes | JSOUCP-MHP | IMD-EACBR | SFO Algorithm | GWO Algorithm | Genetic Algorithm |
| 50 | 0.9875 | 0.9323 | 0.8915 | 0.8640 | 0.8136 |
| 100 | 0.9660 | 0.9313 | 0.8776 | 0.8069 | 0.7964 |
| 150 | 0.9598 | 0.9247 | 0.8578 | 0.7831 | 0.7724 |
| 200 | 0.9562 | 0.9038 | 0.8404 | 0.7701 | 0.7398 |
| 250 | 0.9517 | 0.8955 | 0.7997 | 0.7435 | 0.7031 |
| 300 | 0.9489 | 0.8794 | 0.7833 | 0.7071 | 0.6677 |
| 350 | 0.9405 | 0.8445 | 0.7661 | 0.6900 | 0.6286 |
| 400 | 0.9337 | 0.8307 | 0.7389 | 0.6651 | 0.6032 |
| 450 | 0.9123 | 0.8180 | 0.7275 | 0.6325 | 0.5766 |
| 500 | 0.8910 | 0.8054 | 0.7161 | 0.5999 | 0.5500 |

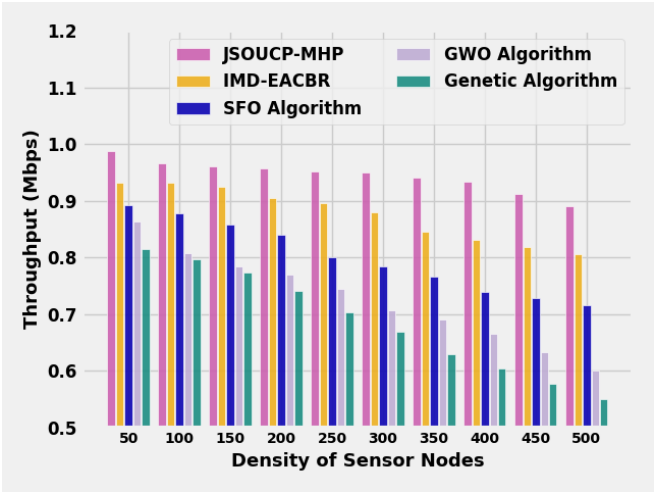



Fig. 6 THROU analysis of JSOUCP-MHP approach under distinct DSNs

Conversely, with 500DSN, the JSOUCP-MHP model has exemplified enhanced THROU of 0.8910Mbps whereas the IMD-EACBR, SFO, GWO, and GA models have depicted reduced NLT of 0.8054Mbps, 0.7161Mbps, 0.5999Mbps, and 0.5500Mbps correspondingly.

A detailed Packet Loss Rate (PLR) valuation of the JSOUCP-MHP with recent methods is exhibited in Tab. 6 and Fig. 7. The results denoted the JSOUCP-MHP approach has resulted in enhanced results with least values of PLR. For example, with 50DSN, the JSOUCP-MHP model has attained least PLR of 0.94% whereas the IMD-EACBR, SFO, GWO, and GA models have gained increased PLR of 1.78%, 3.69%, 3.88%, and 4.96% correspondingly.

TABLE VI

PLR ANALYSIS OF JSOUCP-MHP APPROACH WITH RECENT ALGORITHMS UNDER DISTINCT DSNS

| Packet Loss Rate (%) | | | | | |
|---|---|---|---|---|---|
| Density | JSOUCP- | IMD- | SFO | GWO | Genetic |

| of Sensor Nodes | MHP | EACBR | Algorithm | Algorithm | Algorithm |
|---|---|---|---|---|---|
| 50 | 0.94 | 1.78 | 3.69 | 3.88 | 4.96 |
| 100 | 1.11 | 2.02 | 3.99 | 4.17 | 5.08 |
| 150 | 1.38 | 2.19 | 4.22 | 4.35 | 5.22 |
| 200 | 1.51 | 2.48 | 4.39 | 4.64 | 5.38 |
| 250 | 1.79 | 2.67 | 4.68 | 4.89 | 5.57 |
| 300 | 2.02 | 2.77 | 4.91 | 5.06 | 5.85 |
| 350 | 2.14 | 2.98 | 5.17 | 5.20 | 6.11 |
| 400 | 2.40 | 3.11 | 5.37 | 5.79 | 6.24 |
| 450 | 2.65 | 3.40 | 5.58 | 5.80 | 6.48 |
| 500 | 2.89 | 3.69 | 5.71 | 5.85 | 6.72 |

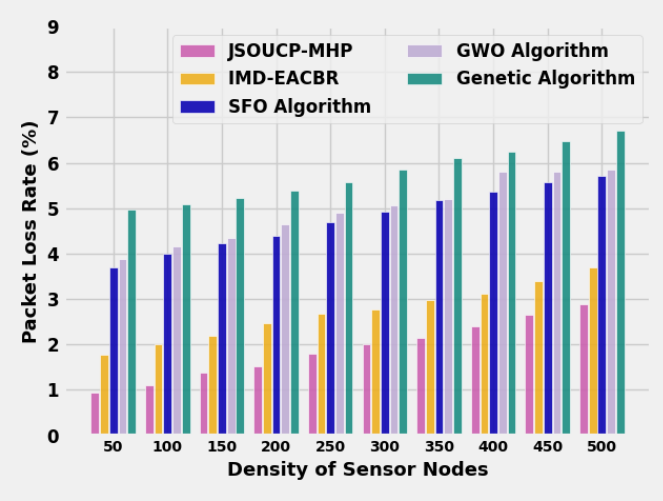



Fig. 7 PLR analysis of JSOUCP-MHP approach under distinct DSNs

Meanwhile, with 250 DSN, the JSOUCP-MHP model has resulted in decreased PLR of 1.79% whereas the IMD-EACBR, SFO, GWO, and GA models have exhibited increased PLR of 2.67%, 4.68%, 4.89%, and 5.57% correspondingly. Finally, with 500 DSN, the JSOUCP-MHP model has offered lesser PLR of 2.89% whereas the IMD-EACBR, SFO, GWO, and GA models have acquired maximum PLR of 3.69%, 5.71%, 5.85%, and 6.72% correspondingly. These results confirmed that the JSOUCP-MHP technique has accomplished enhanced performance over other models.

## IV. CONCLUSIONS

In this article, a novel JSOUCP-MHP system has been presented for resolving hot spot issues in WSN. The JSO algorithm is stimulated by the characteristics of spiders naturally and mathematically modelled the hunting mechanism such as jumping skills, persecution, and search, to attack prey. The presented JSOUCP-MHP technique mainly resolves the hot spot issue for maximizing the network lifetime. To attain this, the JSOUCP-MHP technique elects proper set of CHs using average RE. Also, the JSOUCP-MHP system determines the cluster sizes dependent upon two measures such as RE and DBS. The proposed JSOUCP-MHP technique is examined under several experiments to ensure its supremacy. The comparison study shows the significance of the JSOUCP-MHP technique over other models. In upcoming years, the performance of the JSOUCP-MHP system will be improved by data aggregation approaches.

## REFERENCES

[1] N. Moussa and A. E. B. E. Alaoui, "An energy-efficient cluster-based routing protocol using unequal clustering and improved ACO techniques for WSNs," *Peer to Peer Networking Applications,* vol. 14, no. 3, pp. 1334–1347, 2021.

[2] S. Arjunan and P. Sujatha, "A survey on unequal clustering protocols in wireless sensor networks," *Journal of King Saud University - Computer and Information Sciences,* vol. 33, no. 1, pp. 118, 2021.

[3] A. A. Jasim, M. Y. I. Idris, S. R. B. Azzuhri, N. R. Issa, M. T. Rahman *et al.* "Energy-efficient wireless sensor network with an unequal clustering protocol based on a balanced energy method (EEUCB)," *Sensors,* vol. 21, no. 3, pp. 784, 2021.

[4] K. W. Al-ani, F. B. Abdullah and S. Yossuf, "Unequal clustering in wireless sensor network: a review," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 22, no. 1, pp. 419, 2021.

[5] D. Agrawal and S. Pandey, "Load balanced fuzzy-based unequal clustering for wireless sensor networks assisted Internet of Things," *Engineering Reports,* vol. 2, no. 3, pp. e12130, 2020.

[6] R. Vinodhini and C. Gomathy, "Fuzzy based unequal clustering and context-aware routing based on glow-worm swarm optimization in wireless sensor networks: forest fire detection," *Wireless Personal Communications,* vol. 118, no. 4, pp. 3501–3522, 2021.

[7] S. Phoemphon, C. So-In, P. Aimtongkham and T. G. Nguyen, "An energy-efficient fuzzy-based scheme for unequal multihop clustering in wireless sensor networks," *Journal of Ambient Intelligence and Humanized Computing,* vol. 12, no. 1, pp. 873–2021.

[8] F. Zhu and J. Wei, "An energy-efficient unequal clustering routing protocol for wireless sensor networks," *International Journal of Distributed Sensor Networks,* vol. 15, no. 9, p. 155014771987938, 2019.

[9] N. M. Shagari, M. Y. I. Idris, R. B. Salleh, I. Ahmedy, G. Murtaza *et al.* "A hybridization strategy using equal and unequal clustering schemes to mitigate idle listening for lifetime maximization of wireless sensor network," *Wireless Networks,* vol. 27, no. 4, pp. 2641–2670,

[10] B. M. Sahoo and T. Amgoth, "An improved bat algorithm for unequal clustering in heterogeneous wireless sensor networks," *SN Computer Science,* vol. 2, no. 4, pp. 290, 2021.

[11] K. S. Arikumar, V. Natarajan and S. C. Satapathy, "EELTM: An energy efficient lifetime maximization approach for wsn by pso and fuzzy-based unequal clustering," *Arabian Journal for Science and Engineering,* vol. 45, no. 12, pp. 10245–10260, 2020.

[12] P. S. Mehra, "E-FUCA: enhancement in fuzzy unequal clustering and routing for sustainable wireless sensor network," *Complex & Intelligent Systems*, vol. 8, no. 1, pp. 393–412, 2022.

[13] F. Wang and H. Hu, "An energy-efficient unequal clustering routing algorithm for wireless sensor network," *Revue d'IntelligenceArtificielle*, vol. 33, no. 3, pp. 249–254, 2019.

[14] T. T. Nguyen, J. S. Pan and T. K. Dao, "A compact bat algorithm for unequal clustering in wireless sensor networks," *Applied Sciences,* vol. 9, no. 10, pp. 1973, 2019.

[15] P. C. S. Rao, P. Lalwani, H. Banka and G. S. N. Rao, "Competitive swarm optimization based unequal clustering and routing algorithms (CSO-UCRA) for wireless sensor networks," *Multimedia Tools and Applications,* vol. 80, no. 17, pp. 26093–26119, 2021.

[16] K. Guleria and A. K. Verma, "Meta-heuristic ant colony optimization based unequal clustering for wireless sensor network," *Wireless Personal Communications,* vol. 105, no. 3, pp. 891–911, 2019.

# Deveopment of Accident Prevent Technology for Electric Wheelchair Based on Sensor and Obstacles Notice

Sungmoon Yoo[1], JaeKyung Nam[2], Min Choi[3,] and Seong-A Lee[4]*

[1] *Ministry of Health and Welfare of National Rehabilitation Center, Seoul, South Korea*

[2] *Dept. of Information and Communication Engineering, Chungbuk National University, Cheongju, South Korea*

[3] *Dept. of Information and Communication Engineering, Chungbuk National University, Cheongju, South Korea*

[4]*Department of Occupational Therapy, Soonchunhyang University, Asan, South Korea*

*Contact: kimkim2@sch.ac.kr, myanmy@sch.ac.kr *Corresponding author: Seong-A Lee

Abstract: This study aims to develop a technology that notifies whether obstacles are detected and drive assist wheelchair clients to be aware of the danger to prevent accidents involving electric wheelchairs. As a method, obstacles are recognized by using sensors (ultrasonic waves, radars, etc.) and front cameras as driving assist devices. It devices developed an assistive technology that converts visual signals and information to inform clients of dangerous situations. As a results of the study, There has been developed that the front cameras and ultrasound measurements are used to identify objects in front of them and connects each device in an embedded format though Arduino and MCU boards to announce forward collision warnings within 1m. If you are close to an obstacle through object recognition in front, this device can stop itself and by linking with smartphone applications. It is possible to ask for help from people around it in various ways. Therefore, future research, including backward aims to develop additional devices and identify problems by using them in practice, focusing on devices that are not implemented.

key words: Electric-Wheelchair, Safety, Automatic-limitation.

## I. INTRODUCTION

In Korea, demand for electric wheelchairs is increasing trend of the aging person[1]. It is demand for electric wheelchair is increasing as the aging person increases. According to the Korea Consumer Agency, There is many experienced accidents 35.5% of users of electric security equipment(electric wheelchairs and electric scooters). Accident types include being caught in dumb and obstacles in the street(41.2%), crash with external obstacles(36.3%), crash with vehicles(24.5%), the collision with a pedestrian(22.5%), falling down due to loss of balance of wheelchairs tilt(12.7%), the struck of on electric accessories(10.8%), machine malfunction and fires(8.8%) etc [2]. Personal mobility, such as electric wheelchairs has low posture stability and is exposed to the outside, leading to a risk of conduction or falling down [2]. In this way, to prevent accidents for clients of personal mobile devices such as electric wheelchairs and forward collision-avoidance technology, in this study to be necessary to develop a technology that recognizes obstacles using sensors(ultrasonic/radar, etc), front cameras as driving aids, provides visual signals, alarms to recognize one's own dangerous situation and inform the environment. Therefore, in order to prevent accidents involving electric wheelchairs, this study aims to develop a technology to inform the environment of risks though a driving assistance recognition device so that wheelchair using clients are aware of dangerous

context.

## II. METHODS AND MATERIALS

Electric wheelchairs are designed for environmental awareness, the researcher envisioned to analysis and recognizes the image acquired by connecting the sensor (ultrasonic, front camera and radar, etc) and to use the control device through the embedded device to stop the electric wheelchair if there is an object in front of it within 1m.


Fig. 1 Basic concept of automatic limitation of wheelchair

The manufactured device recognizes and analyses images acquired through sensors(ultrasonic,tadar, etc.) and front cameras in an electric wheelchair to recognize images for clients and provide information to customers. After that, it was conceived to develop a technology for recognizing and controlling front obstacles on embedded boards. It is recognized about data acquired from sensors in personal mobility and images taken from front cameras are recognized through MCU devices and embed. As a result, there was to predict about risk factors in the electric wheelchair environment, if it were full-scale risk factors, an alarm was raised or stopped to pay attention to the elderly or the disabled who are clients [3].

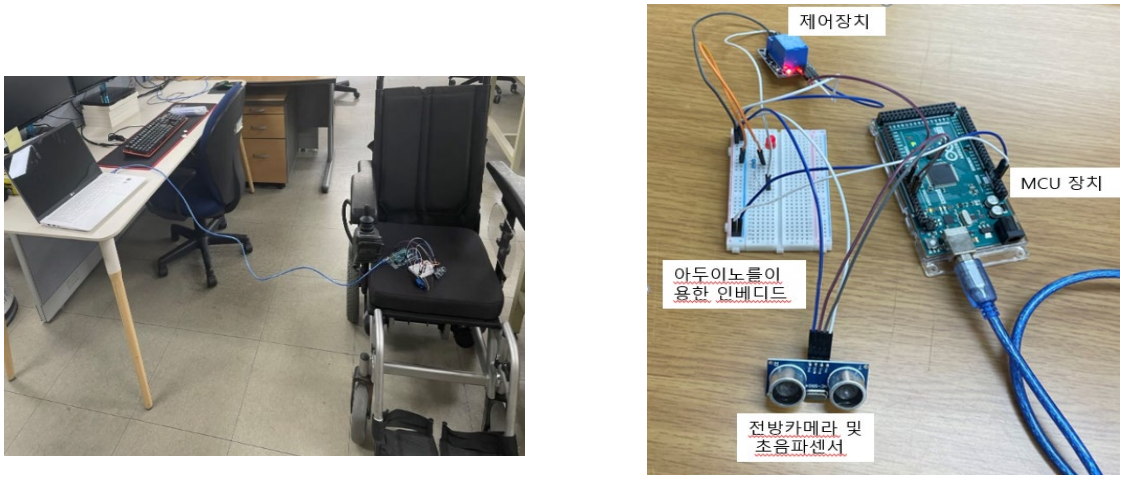



Fig. 2 Experimental setup for automatic limitation of wheelchair

Ultrasonic radar is a non-contact distance measurement sensor widely used to measure distance in cars and HC-SR04 consists of an ultrasonic transmission unit(Trig), an ultrasonic reception unit(echo) and a control circuit. The principle of distance measurement using ultrasonic waves is that the emitted ultrasonic waves hit the object then reflect and return, the time taken at this time is measured and converted into distance. Accordingly, the collected data is tracked in chronological order to calculates the relative speed of the moving object. Also, It calculates the possibility and time of collision with the target vehicle by inferring the distance and direction of movement away from the personal mobile device.

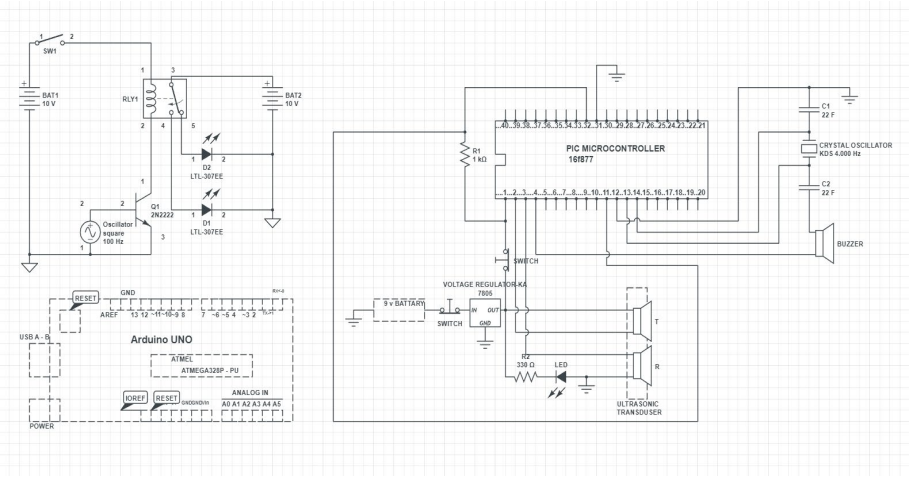



Fig. 3 Schematic diagram of automatic limitation of wheelchair

## III. RESULTS AND DISCUSSION

In this study, there had developed of automatic-limitation of based on the information obtained through sensors and front cameras mounted on personal mobile devices such as electric wheelchairs or personal moving devices, that control device stopped the electric wheelchairs. This research, ICT technology can be applied to electric wheelchairs or a personal mobile device to provide services that will help the elderly and disabled clients who use this device [4, 5]. In addition, clients who elderly and the disabled person will be to reduce electric wheelchair accidents and increase life satisfaction.

## IV. CONCLUSIONS

Through this study, object analysis and reliability data for various objects present in images were provided for data or images acquired from sensor in electric wheelchairs. This data provides the ability to recognize and control forward obstacles on embedded boards. However, by recognizing objects or persons with sensor data acquired from a wheelchair and images taken from a front camera, it is possible to predict many risk factors around it. In addition,

it is necessary to develop additional devices that generate visual signals and warning sounds do that the elderly and the disabled can pay attention when risk factors exist.

## ACKNOWLEDGMENT

This research was supported by Chungbuk National University. And This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2022H1D8A3038040). Also This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) program(IITP-2023-2020-0-01832) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation). This research was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0012724, The Competency Development Program for Industry Specialist) and the Soonchunhyang University Research Fund.

## REFERENCES

[1] Statistics Korea, [online] Available: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1YL20631&conn_path=I 2.

*[2] S.J.Kim, B.K. Kim. Development of Real –Time Control Archiecture for Autonomous Navigation of Powered Wheelchair. Journal of control.automation and systems engineering. Vol10.no 10.* (2004)

*[3]* T.S. Yoon, S.J. Ann, S.M. Kim, Y.B. Han and J.Y. Kim., *Design and Control of Hybrid a Powered Wheelchair for the Elderly*. Trans. Korean Soc. Mech. Eng. Vol. 40, No. 12, pp. 1067~1076, 2016. DOI http://dx.doi.org/10.3795/KSME-A.2016.40.12.1067

*[4]* G.Lawitzky. A Navigation System for Cleanin Robot, Autonomous Robots, Vol. 9, No3, pp255~260, 2000

*[5]* H.G. Kim, D.H.Lee, J. W. Ahn. Design of Safety Drive For Elec Wheelchair Using Ultra Sonic Sensors. Koreasience.org (2023) https://koreascience.or.kr/article/CFKO201135146900388.

# Classification of Optical Coherence Tomography (OCT) Image Using Deep Learning and Ant Colony Optimization

Awais Khan [1,*], Kuntha Pin [1], Jung Woo Han[2], Nab Mat [1], and Yunyoung Nam[3]
[1]*Department of ICT Convergence, Soon Chun Hyang University, Asan, 31538, Republic of Korea*
[2]*Department of Ophthalmology, Bucheon Hospital, Soon Chun Hyang University College of Medicine, Bucheon, 14584, Republic of Korea*
[3]*Department of Computer Science and Engineering, Soon Chun Hyang University, Asan, 31538, Republic of Korea, USA*
*Contact: awaiskhanfa14@gmail.com

*Abstract*— **Optical coherence tomography (OCT) is a widely utilized imaging technique employed in the detection and classification of retinal eye diseases, which are the leading cause of blindness. Nonetheless, manual detection based on OCT images by ophthalmologists can be susceptible to errors and subjectivity. Automated approaches employing deep learning on OCT images are being developed for early detection and classification of retinal diseases, including ARMD, BRVO, CRVO, CSCR, and DME. This research introduces a novel deep learning-based methodology for this purpose. The proposed methodology consists of four fundamental steps. Initially, two pre-trained models, namely InceptionV3, and Resnet50, are modified according to the characteristics of the dataset. Subsequently, transfer learning is employed to extract features, which are further refined and optimized using the Ant Colony Optimization (ACO) algorithm to select the most relevant features. The method utilized K-Nearest Neighbours (KNN) and Support Vector Machine (SVM) algorithms for final classification, achieving a remarkable 99.1% accuracy on OCT retinal images from Soonchunhyang University Bucheon Hospital.**

## I. INTRODUCTION

The retina of the human eye is essential for the processing of visual information. The inner surface of the eyeball is lined by a layer of photosensitive optical nerve tissues [1]. The retina receives light focused by the lens, which produces neural signals. Inside the retina, the macula which is in charge of sharp and sensitive vision perceives visual details, colors, and light intensities. Using the optical nervous system, the retina analyses this data before sending neural impulses to the brain. Eye conditions like macular degeneration, which is often ignored, can cause total [2, 3]. Ophthalmologists typically categorize these types of diseases into five types. Age-related macular degeneration (ARMD), branch retinal vein occlusion (BRVO), central retinal vein occlusion (CRVO), central serous chorioretinopathy (CSCR), and diabetic macular edema (DME) are prevalent conditions that can lead to blindness [4]. These diseases affect the retina, the light-sensitive layer at the back of the eye, and can result in vision loss if left untreated [5].

The availability of supervised data is limited in medical fields, and requires specialized knowledge to create. To overcome this challenge, various deep learning techniques are being developed [6]. One method is to increase the number of training samples by using data augmentation techniques, such as geometrical transformations of images or mimicking image distributions [7, 8]. Another approach is not-so-supervised learning, including semi-supervised, multi-instance, and transfer learning [9]. Transfer learning has gained popularity in recent years, as it effectively transfers model knowledge across different or unrelated tasks, requiring minimal retraining. Kermany et al.[10] demonstrated the effectiveness of transfer learning in a study using deep learning models to classify normal eyes and eyes with three macular diseases, utilizing 4,000 optical coherence tomography images.

To achieve image classification tasks, a number of well-established machine learning (ML) techniques, such as support vector machines (SVM) and random forests, rely on feature extraction methods to generate discriminative representations. For example, Alsaih et al.[11] developed a feature extraction pipeline that integrates the local binary mode of optical coherence tomography (OCT) and the directional gradient histogram to construct a distinctive feature set. This pipeline uses a linear SVM classifier to predict image categories. Similarly, Sun et al.[12] introduced a universal approach for retinal image alignment and cropping, which is followed by multiclass linear SVM classification to categorize dry age-related macular degeneration (AMD) and diabetic macular edema (DME). The global image representation is obtained through sparse coding and a spatial pyramid[13].

The goal of this study was to overcome the limitations of existing approaches by proposing a new deep learning and Ant Colony Optimization framework for precise OCT image classification. The proposed framework includes the following steps:

- Modification of three pre-trained models, Inception-V3 and Resnet-50, by adding a new layer that connects the preceding layers in terms of fully connected layers.
- The ant colony optimization (ACO) method is used for the feature selection. This method involves first selecting features with ACO, which are then fine-tuned with an activation function.
- The ACO was applied to both modified deep learning models to compare accuracy. Based on accuracy, the best one is selected for the final classification.

## II. Materials and Methods

### A. Dataset

The OCT (Optical Coherence Tomography) images used in this study were obtained from the Soon Chun Hyang University Bucheon Hospital, and were labelled with eye diseases by ophthalmologists at the hospital. The images were collected and normalized with approval from the Institutional Review Board (IRB). The dataset was collected twice, with the first collection consisting of 2000 images in 2021 and the second collection consisting of 998 images in 2022. This study was approved by the IRB of Soon Chun Hyang University Bucheon Hospital in Buchoen, Republic of Korea (IRB approval number 2021-05-001).

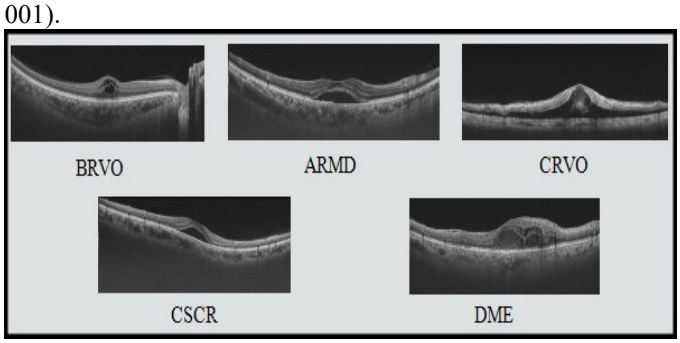



Figure 1. Sample of OCT image dataset.

### B. Proposed Methodology

For the OCT image classification, we propose a new deep learning method, which is represented in the main flow diagram in Figure. 2. The methodology consists of a sequence of several steps, namely data preprocessing, feature extraction employing pretrained models, feature optimization, and lastly, classification. This strategy implements advanced deep transfer learning techniques with the objective of enhancing two pretrained models, specifically Resnet 50 and Inception V3 [36]. Upon extraction of the features from these modified models, the resulting vectors are subsequently refined through the utilization of an advanced algorithm known as ant colony optimization. Lastly, the refined features are subjected to multiclass classification methods to obtain the final results.

### C. ResNet-50 and Inception-V3

The neural network architecture handles image classification tasks with 48 layers and is trained on 1000 object classes. The input size for images is 299x299x3. It includes convolutional layers, max pooling, normal pooling, concatenation, dropouts, and fully connected layers. The network passes the input image through three convolutional layers and a max pooling layer. This modified model, trained on OCT image dataset, has a constant input size of 224x224x3. Transfer learning techniques were applied to train the model using features from the average pooling layer, resulting in a feature vector of size N×1920.

ResNet architecture solves the vanishing gradient problem and enhances information flow during training. The input size for ResNet-50 is 224x224. With 23 million parameters, we customized ResNet-50 for our OCT image dataset by adding a new fully connected layer. Through transfer learning, we extracted features, resulting in N × 2048dimensional feature vectors.

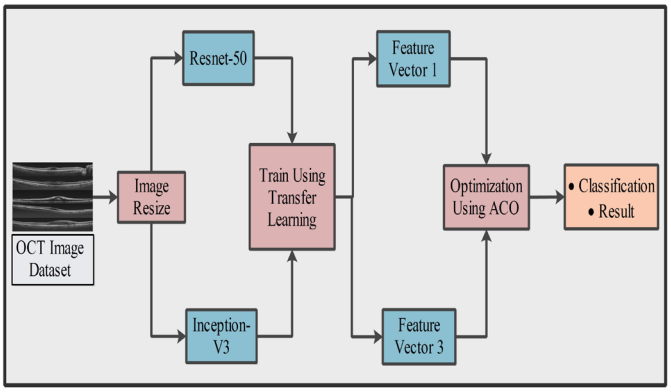



Figure 2. Main flow diagram of the proposed method, features are extracted by transfer learning, make a feature vector, ACO is applied on the feature vector for optimization and at last classification.

## III. Results and Discussion

The In this section, OCT images are utilizes for the experiment process. The modified deep learning models named ResNet50 and InceptionV3 were employed to generate the results. The performance of the ResNet50 model is presented in Table I. The results indicate that the Cubic support vector machine (CSVM) performed well as compared to the other proposed machine learning (ML) classifiers. The highest accuracy using the ResNet50 was 97.3% using the Cubic SVM. The precision rate, recall rate and AUC were 98%, 97.6% and 1. The Quadratic SVM got second best accuracy of 97.6% with the precision rate, recall rate and AUC of 97.1%, 96.1% and 1. The Linear SVM and Cosine KNN the accuracies were 95.4% and 94.3% respectively. The accuracy difference in between the Quadratic SVM and Cubic SVM was approximately 0.4%. The confusion matrix for ResNet50 model and ACO is presented in Figure 3.

Table II. present the classification results of the InceptionV3 model. The optimized features are passed to four ML classifiers; the results shows that the Cubic SVM perform well as compared to the other classifiers and obtain the accuracy of 96.3%. The precision rate, recall rate and the AUC were 96.1%, 93.6% and 1. The Quadratic SVM achieved the second-best accuracy of 95.9%, with precision rates, recall rates, and AUC of 97.1%, 96.1%, and 1. The third best accuracy was 93.9% achieved by the Linear SVM with the precision rate, recall rate,

and AUC of 92.8%, 90.14%, and 0.99.

Table I. Proposed classification result of OCT image using 10-fold cross validation, ResNet50 and ACO.

| **Classifiers** | **Accuracy** | **AUC** | **Precision Rate** | **Recall Rate** |
|---|---|---|---|---|
| **Cubic SVM** | **97.3%** | **1** | **98** | **97.6** |
| Quadratic SVM | 97.6% | 1 | 97.1 | 96.1 |
| Linear SVM | 95.4% | 0.99 | 95 | 92.2 |
| Cosine KNN | 94.3% | 0.99 | 94.8 | 91.4 |

Table II. Proposed classification result of OCT image using 10-fold cross validation, InceptionV3 and ACO.

| Classifiers | Accuracy | AUC | Precision Rate | Recall Rate |
|---|---|---|---|---|
| **Cubic SVM** | **96.3%** | **1** | **96.1** | **93.6** |
| Quadratic SVM | 95.9% | 1 | 95.3 | 93.4 |
| Linear SVM | 93.9% | 0.99 | 92.8 | 90.14 |
| Cosine KNN | 90.4% | 0.99 | 90.3 | 86.2 |

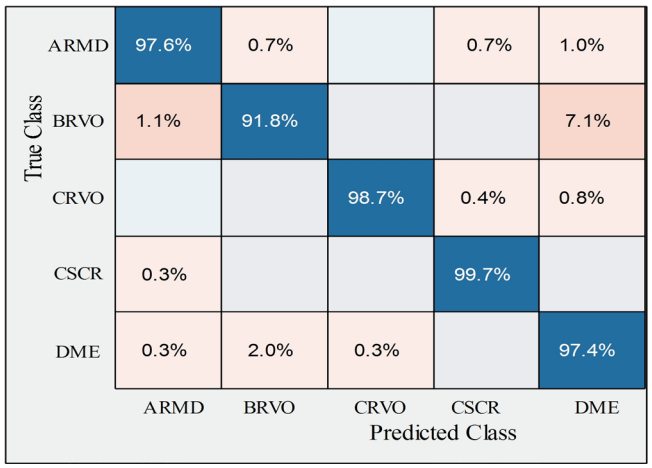



Figure 3. Confusion matrix of cubic SVM using the modified ResNet50.

## IV. Conclusions

This study proposed a novel method for the detection of retinal eye diseases using retinal OCT images. The results of the proposed method showed that the combination of deep features and the proposed hybrid selection process achieved the best performance with an average accuracy of 99.1%. A comparative study with existing approaches was also performed, and the results showed that the proposed method outperformed the other methods with the highest accuracy. It can be concluded that the proposed method is reliable and can accurately classify retinal eye diseases. In the future, the method can be further tested and evaluated on different datasets by adding more eye diseases and different feature selection methods. The potential impact of this method is significant as it can aid in early and accurate diagnosis of retinal eye diseases, leading to prompt and effective treatment.

## Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(No. NRF-2021R1A2C1010362) and the Soonchunhyang University Research Fund.

## References

[1] M. J. Umer, M. Sharif, M. Raza, and S. J. E. S. Kadry, "A deep feature fusion and selection-based retinal eye disease detection from OCT images," *Expert Systems,* p. e13232.

[2] A. Yuan and P. K. Kaiser, "Emerging therapies for the treatment of neovascular age related macular degeneration," in *Seminars in ophthalmology*, 2011, vol. 26, no. 3, pp. 149-155: Taylor & Francis.

[3] D. S. Kermany *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *cell,* vol. 172, no. 5, pp. 1122-1131, 2018.

[4] Ş. Ţălu and S. D. J. W. J. o. D. Nicoara, "Malfunction of outer retinal barrier and choroid in the occurrence and progression of diabetic macular edema," *World Journal of Diabetes,* vol. 12, no. 4, p. 437, 2021.

[5] Z. Ai *et al.*, "FN-OCT: Disease detection algorithm for retinal optical coherence tomography based on a fusion network," *Frontiers in Neuroinformatics,* vol. 16, 2022.

[6] G. An, M. Akiba, K. Omodaka, T. Nakazawa, and H. J. S. r. Yokota, "Hierarchical deep learning models using transfer learning for disease detection and classification based on small number of medical images," *Scientific reports,* vol. 11, no. 1, p. 4250, 2021.

[7] J. Ker, L. Wang, J. Rao, and T. J. I. A. Lim, "Deep learning applications in medical image analysis," *Ieee Access,* vol. 6, pp. 9375-9389, 2017.

[8] I. Goodfellow *et al.*, "Generative adversarial networks," *Communications of the ACM,* vol. 63, no. 11, pp. 139-144, 2020.

[9] V. Cheplygina, M. de Bruijne, and J. P. J. M. i. a. Pluim, "Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis," *Medical image analysis,* vol. 54, pp. 280-296, 2019.

[10] L. Torrey and J. J. I. G. Shavlik, "Transfer learning. Handbook of research on machine learning applications," *IGI Global,* vol. 3, pp. 17-35, 2009.

[11] K. Alsaih *et al.*, "Classification of SD-OCT volumes with multi pyramids, LBP and HOG descriptors: application to DME detections," in *2016 38th Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, 2016, pp. 1344-1347: IEEE.

[12] Y. Sun, S. Li, and Z. J. J. o. b. o. Sun, "Fully automated macular pathology detection in retina optical coherence tomography images using sparse coding and dictionary learning," *Journal of biomedical optics,* vol. 22, no. 1, pp. 016012-016012, 2017.

[13] Z. Ai *et al.*, "FN-OCT: Disease detection algorithm for retinal optical coherence tomography based on a fusion network," vol. 16, 2022.

# Classification of hospitalized patients at risk of falling based on machine learning and long-term movement monitoring data

Jung-Yeon Kim[1,*], Chomyong Kim[1], Seob Jeon[2], Hyo-Wook Gil[2], Ik-Dong Yu[2], Ji-Won Lyu[2], Euy-Hyun Chung[2], and Yunyoung Nam[3]

[1]*ICT Convergence Research Centre, Soonchunhyang University, Asan, South Korea*

[2]*Soonchunhyang University Cheonan Hospital, Cheonan, South Korea*

[3]*Department of Computer Science and Engineering, Soonchunhyang University, Asan, South Korea*

*Contact: betterwithme@sch.ac.kr

*Abstract*— **Fall risk assessment has been an active research topic since falls are the leading cause of injury-related death among older adults. However, Majority number of studies seem to focus on short-term sensing data collected while evaluating gait and balance using clinical assessment tools. Investigating only the short-term monitoring data may be insufficient as it is possible that fall related symptoms are not presented during fall risk assessments. Thus, collection and analysis of long-term gait monitoring is necessary although this poses some problem to address, such as removing irrelevant data that do not contain clinically meaningful data. In this study, we describe the development processes of three core modules and report results of performance on gait recognition and classification for patients at risk of falls. Each module is responsible for separating gait patterns from irrelevant data, computing parameters associated with gait analysis, and finally classifying hospitalized patients with high risk of falls, respectively. The result of testing the proposed method suggest that machine learning based approaches can help reduce the data volumes of long-term monitoring data and identify patients at high risk of falls with the confidence.**

## I. Introduction

It is known that falls are not only the leading cause of death among adults aged 65 and over, but also reduce the quality of life. Moreover, people with reduced functioning, for example, hospitalised patients, can be severely affected by falls. Owing to this, it has been an active research topic to analyse risk factors of falls in older adults [1]. Although considerable number of studies conducted to investigate ways of utilizing advanced sensor technologies and machine learning techniques in the field [2-6], it appears that majority number of studies focus on short term sensing data, such as inertial sensors, collected during conducting clinical assessment tools for gait and balance [2, 7]. While inertial sensors can be beneficial to capture long-term movement data as they can be easily worn, analysing long-term monitoring data can be problematic as long-term monitoring data collect high volumes of data due to the relatively high sampling rate set to collect gait pattern. Not only does the large amount of data require high computational resources, but also mass storage, which become problematic when it comes to scalability as the amount of data collected by sensors increase exponentially. In addition, long-term movement data include non-gait movements that contain little information about gait quality because most inpatients are relatively inactive that the data collected for a long period of time contain little information, such as lying down on a bed or sitting on a sofa. To address the issues, it is crucial to only identify gait from the data and keep them for analysis.

Long-term data must be processed to ensure that the data only include gait patterns and collect clinically meaningful information, excluding irrelevant data. Not only does this process help to save computation resource, but also storage space. In this study, we aimed to describe ways of separating gait patterns from irrelevant data, computing parameters associated with gait analysis, and finally classifying hospitalised patients with high risk of falls. More detailed information is provided in the next chapter.

## II. Methods and Materials

### A. Data collection

Two different kinds of datasets were collected for developing gait recognition module and classification module for identifying patients at high risk of fall. For the development of gait recognition module, healthy participants (n=24) were recruited to simulate movements such as walking, sitting, and lying, and inertial data were collected by IMU sensor worn on their both ankles. While the simulated data collected during walking activity was labelled as gait and others were labelled as non-gait. On the other hand, hospitalized patients (n=28) were recruited for data collection so that we can test gait recognition module and develop a fall risk level classification model. The patients were only asked to wear the same IMU sensor on their both ankles. No other instructions were given to the patients during the data collection process. When patients are admitted to the hospital, degree of fall risk is assessed by nurses using More Fall Scale (MFS), which can be used to categorize patients into three groups: high, moderate, and low risk of fall. The result of MFS was used to label the dataset so that they can be used to develop classification model that identifies patients at risk of high risk of fall.

### B. System overview

Machine learning algorithms were used to implement gait recognition module and classification module for identifying patients at high risk of falls whereas rule-based approach was

applied to identify a pair of heel-strike and toe off, which compose a gait cycle, and gait related parameters were computed. The gait recognition model was developed by training and testing machine learning algorithms on the data collected over a long period of time and labelled to either "gait" or "non-gait". After gait recognition process, gait parameters were computed from the processed data that only contain gait patterns, and used to develop the classification model that aims to differentiate patients whose higher risk of falling, which had been categorized according to Morse Fall Scale (MFS) of individual patients. The result of machine learning based gait pattern recognition performance and classification of patients with high risk of fall will be reported.

### C. Gait recognition module

This module aims to process long-term monitoring data and keep parts of data that are only necessary for fall risk assessment as well as gait analysis. To achieve this, the module employs machine learning techniques that learns characteristics of features reflecting gait or non-gait patterns and classify them accordingly. The machine learning algorithm tested in this study include random forest (RF), support vector machine (SVM), multi-layer perceptron (MLP), Adaboost (ADA) and XGBoost (XGB). The datasets were pre-processed to remove influences of gravity and motion artifacts caused by sensor displacement. Following this, statistical features were computed from segments with 50% overlapping obtained by sliding a window with the length of 2.5 seconds. For the gait recognition model development, 11,114 segments that were labelled either gait or non-gait, but the number of non-gait data samples were relatively less than gait data samples. Totally, 12 statistical features were computed including mean, standard deviation, root mean squared value, maximum and minimum amplitudes, median, skewness, kurtosis, first and third quartile, auto-correlation, and zero crossing counts. To address data imbalance, feature sets that were labelled as gait were augmented after extracting features from the segments using the synthetic minority over-sampling technique, which resulted in 18,549 segments balanced in total. K-fold cross validation was applied during the model development, and data augmentation was only done for the hold out data used to train the model. Next, the feature sets were fed to machine learning algorithms to identify gait from the entire data. Gait recognition model was trained and tested with the dataset containing movements simulated by healthy participants. Finally, the gait recognition model was used to only keep the gait parts of data collected from hospitalised patients over a long period of time.

### D. Module for computation of gait parameters

After the gait recognition module identified only gait signals, this module determined heel-strike and toe-off events according to series of condition rules to differentiate them by analysing the gait signals in time domain. Then, gait parameters in both time domain and frequency domain were computed resulting in as summarized in Table 1.

### E. Module for classification of patients at high risk of fall

While time domain gait parameters are generally used for statistical analysis to obtain insight on individual gait quality, they were fed into the classification model to test if machine learning algorithms could identify patients at high risk of fall in this study. Machine learning algorithms tested for this model are the same ones used to develop the gait recognition module. The number of recruited inpatients is small that leave-one-out cross-validation scheme was applied to test classification performance.

TABLE I. LIST OF PARAMETERS COMPUTED IN THIS STUDY

| Types of parameters | | Parameters |
|---|---|---|
| Degree of active state | | Percentage of gait parts from the data |
| Time domain in dataset of both legs | swing | Mean, standard deviation, variance, median, IQR, ratio of swing or stance phases to stride |
| | stance | |
| | stride | Mean, standard deviation, variance, median, IQR |
| | ratio | Absolute value of (1 - (gait params from left leg) / (gait params from right leg)) |
| Frequency domain in dataset of both legs | | Dominant frequency, power, energy |

### F. Analysis

Performance of gait recognition and classification was evaluated using sensitivity, specificity, accuracy, and F1-score according to machine learning algorithms. On the other hand, statistical analysis was conducted on gait parameters. Statistical similarity of the parameters between three groups (high, moderate, and low risk group) of patients determined by the MFS score was tested using one way analysis of variance and Bonferroni test for post-hoc analysis. SPSS (V.23) Window version was used for the statistical analysis and significance level ($\alpha$) was set at 0.05.

## III. RESULTS

Performance of gait recognition and classification of patients at risk of fall and statistical analysis on gait parameters according to MFS score categories are reported here.

### A. Gait recognition performance

As shown in Table 2, gait recognition performance was over 99% for machine learning algorithms tested in this study with the standard deviation ranging between 1 and 5%. Although classification accuracy of RF was the highest, the difference of accuracy between the tested models was marginal. The result indicates that machine learning approach can successfully determine gait from other types of movements or non-movement state.

TABLE III. GAIT RECOGNITION PERFORMANCE

| | Sensitivity (%) | Specificity (%) | Accuracy (%) | F1-score (%) |
|---|---|---|---|---|
| RF | 99.9 | 99.8 | 99.9 | 99.9 |
| SVM | 99.8 | 98.9 | 99.4 | 99.4 |
| MLP | 99.8 | 99.2 | 99.5 | 99.5 |
| AdaBoost | 99.9 | 99.8 | 99.9 | 99.9 |
| XGBoost | 99.9 | 99.8 | 99.9 | 99.9 |

### B. Gait characteristics of patients

Statistical analysis of gait parameters was compared between groups using one-way ANOVA, and those gait parameters with

TABLE IIIII. GAIT PARAMETERS COMPUTED FROM IDENTIFIED PAIRS OF HEEL STRIKE AND TOE OFF

| Parameters | Low | Moderate | High | F (p) |
|---|---|---|---|---|
| Ratio of swing phase of left leg | 0.501 | 0492 | 0.462 | 0.372 (0.040) |
| Ratio of stance phase of left leg | 0.498 | 0507 | 0.537 | 0.372 (0.040) |
| Ratio of swing phase of right | 0.529 | 0.487 | 0.468 | 7.877 (0.001) |
| Mean stance phase of right leg | 0.351 | 0.460 | 0.472 | 5.167 (0.008) |
| Std of stance phase of right leg | 0.328 | 0.366 | 0.396 | 7.134 (0.002) |
| Variance of stance phase of right leg | 0.109 | 0.139 | 0.160 | 7.702 (0.001) |
| Median of stance phase of right left | 0.234 | 0.357 | 0.341 | 4.441 (0.015) |
| IQR of stance phase of right leg | 0.302 | 0.413 | 0.413 | 4.123 (0.020) |
| Ratio of stance phase of right leg | 0.470 | 0.512 | 0.531 | 7.877 (0.001) |
| Mean stride of right leg | 0.739 | 0.877 | 0.877 | 3.922 (0.024) |
| Std stride of right leg | 0.574 | 0.610 | 0.639 | 4.143 (0.020) |
| Variance stride of right leg | 0.170 | 0.199 | 0.219 | 5.545 (0.006) |
| Median stride of right leg | 0.544 | 0.714 | 0.678 | 3.958 (0.024) |
| IQR stride of right leg | 0.527 | 0.661 | 0.658 | 3.188 (0.047) |
| Ratio of power of left to right leg | 0.176 | 0.370 | 0.220 | 3.193 (0.047) |
| Ratio of power of left to right leg | 0.176 | 0.370 | 0.220 | 3.193 (0.047) |

statistical difference are summarized in Table III. Although 16 gait parameters were found to be statistically different between groups, post-hoc analysis indicates that the significant differences were only found either between low-risk group and high-risk group or between low-risk group and moderate-risk group, but not between moderate-risk group and high-risk group. This suggests that the use of MFS to categorize patients into high, moderate, and low risk group may not be realized by the gait parameters computed by the proposed method.

### *C. Classification of patients at high risk of fall*

Classification model was designed to categorize patients into three groups: high, moderate, and low risk of fall. However, the result of classification of patients into three different risk groups were relatively poor ranging from 0.61 to 0.77 as shown in Table IV. Due to the relatively poor classification performance into three groups, the groups were reduced to two groups by merging moderate-risk group and high-risk group into high-risk group. Table V summarize classification performance for the two groups. The highest accuracy was achieved with MLP classifier (0.95), and the rest models shows the accuracy ranging from 0.91 to 0.92.

## IV. CONCLUSIONS

This study described how the data were collected to build gait recognition model and classification model to identify patients at risk of fall. Both healthy adults and hospitalized patients were recruited for the data collection, and machine learning algorithms were utilized to identify gait from long-term monitoring data and patients at high risk of fall using gait parameters from the identified gait. While the machine learning based method was successful for gait recognition, it did not sufficiently categorize patients into low, moderate, and high risk groups. Nevertheless, the approach sufficiently categorized fall risk groups into two degrees with up to 95% accuracy. Since the number of datasets used for classifying the degree of patients' risk of fall was collected from less than 30 patients, it still requires more datasets for further research. Even though findings we observed in this study is only limited to some extent, it seems feasible to realize the efficacy of the machine learning based method for fall risk assessment.

TABLE IVV. FALL RISK DEGREE CLASSIFICATION INTO HIGH, MODERATE, AND LOW RISK GROUPS

| | Precision (%) | Recall (%) | Accuracy (%) |
|---|---|---|---|
| RF | 0.72 | 0.73 | 0.73 |
| SVM | 0.76 | 0.77 | 0.77 |
| MLP | 0.73 | 0.72 | 0.72 |
| AdaBoost | 0.64 | 0.61 | 0.61 |
| XGBoost | 0.71 | 0.71 | 0.71 |

TABLE V
FALL RISK DEGREE CLASSIFICATION INTO HIGH AND LOW RISK GROUPS

| | Precision (%) | Recall (%) | Accuracy (%) |
|---|---|---|---|
| RF | 0.91 | 0.91 | 0.91 |
| SVM | 0.93 | 0.92 | 0.92 |
| MLP | 0.95 | 0.95 | 0.95 |
| AdaBoost | 0.91 | 0.91 | 0.91 |
| XGBoost | 0.92 | 0.92 | 0.92 |

## ACKNOWLEDGMENT

This research was supported by Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (P0012724, The Competency Development Program for Industry Specialist).

## REFERENCES

[1] H. Nagano, and R. K. Begg, "Shoe-Insole Technology for Injury Prevention in Walking," *Sensors,* vol. 18, no. 5, pp. 1468, 2018.

[2] S. K. Byrnes, C. Nüesch, S. Loske, A. Leuenberger, S. Schären, C. Netzer, and A. Mündermann, "Inertial Sensor-Based Gait and Attractor Analysis as Clinical Measurement Tool: Functionality and Sensitivity in Healthy Subjects and Patients With Symptomatic Lumbar Spinal Stenosis," *Frontiers in Physiology,* vol. 9, 2018-August-14, 2018.

[3] D. Rodríguez-Martín, C. Pérez-López, A. Samà, J. Cabestany, and A. Català, "A wearable inertial measurement unit for long-term monitoring in the dependency care area," *Sensors,* vol. 13, no. 10, pp. 14079-14104, 2013.

[4] A. Mehmood, M. A. Khan, M. Sharif, S. A. Khan, M. Shaheen, T. Saba, N. Riaz, and I. Ashraf, "Prosperous Human Gait Recognition: an end-to-end system based on pre-trained CNN features selection," *Multimedia Tools and Applications*, 2020/04/28, 2020.

[5] C. Tunca, G. Salur, and C. Ersoy, "Deep learning for fall risk assessment with inertial sensors: Utilizing domain

knowledge in spatio-temporal gait parameters," *IEEE journal of biomedical and health informatics,* vol. 24, no. 7, pp. 1994-2005, 2019.

[6] A. Sampath Dakshina Murthy, T. Karthikeyan, and R. Vinoth Kanna, "Gait-based person fall prediction using deep learning approach," *Soft Computing*, pp. 1-9, 2021.

[7] J.-Y. Kim, S. Lee, H. B. Lee, B.-G. Kang, S.-B. Im, and Y. Nam, "Gait analysis in patients with neurological disorders using ankle-worn accelerometers," *The Journal of Supercomputing,* vol. 77, no. 8, pp. 8374-8390, 2021/08/01, 2021.

# Energy-Efficient Clustering Using Optimization with Locust Game Theory

*P.Kavitha Rani[1], Hee-Kwon Chae[2], Yunyoung Nam[2,*] and Mohamed Abouhawwash[3,4]*

*[1]Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore-641008, India.*
*[2]Department of ICT Convergence, Soonchunhyang University, Asan, 31538, Korea.*
*[3]Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.*
*[4]Department of Computational Mathematics, Science, and Engineering (CMSE), Michigan State University, East Lansing, MI, 48824 USA.*
**Corresponding Author: Yunyoung Nam. Email: ynam@sch.ac.kr*

*Abstract*— **Wireless sensor networks (WSNs) are made up of several sensors located in a specific area and powered by a finite amount of energy to gather environmental data. WSNs use sensor nodes (SNs) to collect and transmit data. However, the power supplied by the sensor network is restricted. Thus, SNs must store energy as often as to extend the lifespan of the network. In the proposed study, effective clustering and longer network lifetimes are achieved using multi-swarm optimization (MSO) and game theory based on locust search (LS-II). In this research, MSO is used to improve the optimum routing, while the LS-II approach is employed to specify the number of cluster heads (CHs) and select the best ones. After the CHs are identified, the other sensor components are allocated to the closest CHs to them. A game theory-based energy-efficient clustering approach is applied to WSNs. Here each SN is considered a player in the game. The SN can implement beneficial methods for itself depending on the length of the idle listening time in the active phase and then determine to choose whether or not to rest. The proposed multi-swarm with energy-efficient game theory on locust search (MSGE-LS) efficiently selects CHs, minimizes energy consumption, and improves the lifetime of networks. The findings of this study indicate that the proposed MSGE-LS is an effective method because its result proves that it increases the number of clusters, average energy consumption, lifespan extension, reduction in average packet loss, and end-to-end delay.**

**Keywords:** Wireless sensor network; clustering; routing; cluster head; energy consumption; network's lifetime; multi swarm optimization; game theory.

## I. INTRODUCTION

Wireless Sensor Networks (WSNs) are a collection of cost-effective, multipurpose Sensor Nodes (SNs) which work together to detect a specific area as an Area of Interest (AoI). The system collects the data from the AoI and sends it to a Base Station (BS) for processing. WSNs have made it simpler to monitor remote places. They are particularly successful in gathering information in a variety of remote locations, including rescue teams, forests, war-prone regions, undersea research, and weather patterns. WSNs have numerous SNs that are connected to a BS. However, it is difficult to power and replace SNs in remote places [1].

SNs in WSNs help to detect communication and gather data collaboratively. A certain quantity of energy is consumed during this procedure. However, SNs are powered by a rechargeable battery and have a finite quantity of energy in case the batteries die. Also, they do not need to change the battery pack or update it promptly. The communication of the data collected will be hampered to some extent, and the entire sensor network may become incapacitated [2]. Consequently, extending the lifespan of the complete sensor network is energy-efficient and has become a major challenge in real-world remote sensing applications.

In real-world applications, wireless SNs are powered by limited-capacity batteries and WSNs.This node is placed unprotected or in unsafe places in a challenging way. They are powered by batteries, and the surveillance region is unregulated. This makes it difficult to store energy supplies and rechargeable batteries. Thus, decreasing the energy consumption of the network seems to be the most effective method to extend the lifespan of the sensor network [3] and ensure the continuous operation of WSNs. Many professionals and academicians have undertaken various types of research to optimize network energy consumption to a certain extent [4–9].

Cluster-based or hierarchical routing is a common mode of communication with appealing qualities such as efficiency and flexibility. The general concepts of routing protocol have been combined with most of the energy-saving approaches in WSNs. Nodes with lots of energy are evaluated for their abilities to analyze and communicate the data in the framework of a hierarchical technique [10]. Low-energy terminals on either side are utilized to gather data in areas close to the objective. Hierarchical navigation is an effective method to reduce power usage within clustered architectures by aggregating the data. Hierarchical routing approaches combine operations to minimize the number of transmitted packets for delivery to a specific sink [11].

In these circumstances, assigning special roles to nodes can considerably extend the lifespan of the network and the flexibility of design. Cluster Heads (CHs) are special nodes that serve as anchors in the hierarchy or cluster-based systems.

Scaling is a key attribute in WSNs. Owing to the limits given by the original assumption, this idea is still a challenge that hasn't been solved in most of the routing approaches. Cluster-based methods find single sinks and a few CHs in a general WSN to increase the influence of their coverage spaces [12].

The assignment of unique purposes to nodes in such circumstances can considerably extend the lifetime of the system and the adaptability of technology. CHs are the special nodes that serve as referrals in hierarchical or clustered systems. The Locust Search (LS-II) method [13] is a meta-heuristic optimization technique derived from the modeling of desert locust biological behavior. A locust acts in two ways; alone and in groups. The solitary method prevents the components from interacting with other entities to discover new sources of food. The sociological perspective, on either hand, takes into account the large proportion of entities in locations with ample food supplies.

The locust concentrating process includes the movement of an entity in lowland areas to all those components in locations with the greatest food supplies. The LS-II technique combines both tendencies to produce highly powerful global and regional searching capabilities.

These properties have led to the application of LS-II in a wide range of challenging optimization formulas, including parameterization in chaotic systems [14], image processing [15–16], and pattern recognition [17].

The main contributions of the proposed method are given below:

- identify the CHs effectively and minimize energy consumption using the LS-II algorithm
- Multi-Swarm Optimization (MSO) helps to find the optimal routing path, and the SN energy is stored feasibly.
- A penalty technique is designed to motivate SNs to embrace collaborative tactics in operational plans to prevent their destructive behavior if they go to sleep.
- For sensor networks exhibiting selfish behavior, the optimum amount of penalty sessions are established to efficiently improve the lifetime of the network.

The remaining sections of this paper are structured as follows; Section 2 discusses the related research. Section 3 describes the WSNs and optimization methods for minimizing energy consumption. Section 4 discusses the experimental results and comparison. Section 5concludes the proposed optimization method with future work.

## II. Related work

Whenever an unbalanced situation occurs in information transfer across the transmission network, an energetic vacuum is generated, causing the sink nodes to die prematurely and reducing the lifespan of the network [18,19]. Researchers [20,21] offered a strategy wherein the communication range was changed depending on the proximity between the CHs and their users. The use of the firefly optimization technique also extended the longevity of the program. The method was applied against the standardized methods in a variety of scenarios. According to the findings of their research, the proposed method produced good outcomes in terms of network life. The author, on the other hand, made no mention of the energy maximization of SNs while preserving the optimal CH range.

The researcher [22] proposed the particle swarm optimization-based unequal fault-tolerant clustering (PSO-UFC) technology. Unbalanced grouping and fault-tolerant difficulties of the present energy-balanced unequal clustering technique were solved effectively in the method for the long-term network. An absence of inequity in the clustering method was used to equalize intra-cluster and inter-cluster power use among the master CHs (MCHs) to find a remedy for the unbalanced in the clustering algorithm. Moreover, with the PSO-UFC method, the internet connection was restored by selecting an additional CH, termed substitute CH, in the event of an impetuous MCH breakdown. The lifespan maximization of SNs, on the other hand, has appeared to be dependent on an optimized clustering method that reduces end-to-end delay.

In [23], the author proposed regional energy-aware clustering with isolated nodes (REAC-IN) for grouping. The CH was chosen in REAC-IN based on the weight, and the value was calculated using the residual energy of each sensor and the average overall energy of all sensors in each cluster. In the end, the proposed approach had good results when compared to the other WSN protocols, although numerous difficulties had been overlooked. A fresh, enhanced method was developed to overcome those limitations. Grouping and task selection were used to demonstrate the suggested technique which could transcend the limits of the REAC-IN routing algorithm. Their suggested methodology surpassed other methodological approaches.

The researchers [24] used an upgraded harmony search (HS)-based routing to transport the incoming packets between CHs and the sinks from another study. The efficiency of the combined grouping and forwarding system was assessed using energy cost, the proportion of live and dead components, and system longevity. In comparison with the existing methods, the novel cuckoo–HS-based combined navigation and grouping method provided more results. When constructing the state-of-the-art meta-heuristic algorithms with merged forwarding and clustering, the researcher disregarded incorporating the lifespan and end-to-end latency elements.

## III. Proposed MSGE-LS Methodology

The proposed multi-swarm with energy-efficient game theory on locust search (MSGE-LS) method consists of two models; the energy model and the network model. The proposed method efficiently selects the CH and optimally chooses the route for routing. It effectively minimizes energy consumption and improves the life span of the network. Fig.1 shows the architecture of the proposed method. MSO with game theory optimally chooses the routing path, and LS-II helps to select the CH successfully.

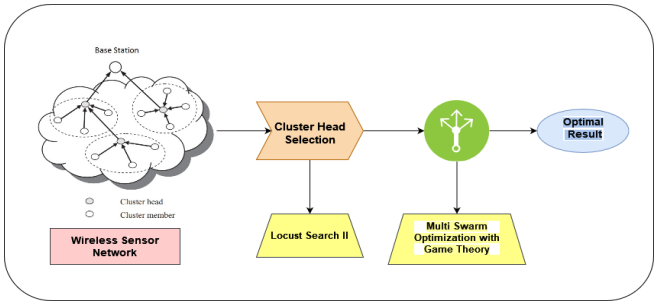



Figure. 1 Architecture of the proposed method

### A. Network Model

A WSN is a network made up of a randomized number of nodes that form various groups, each with a CH node and many cluster members (CMs). Every CM has a detection range through which each node can receive information from the observed item and transmit it to the appropriate CH, who subsequently sends it to the sink node.

The functioning phases of an SN (exit phase) occur when the node has to transmit and receive information, and the inactive monitoring phase occurs when a network has no information to transmit and receive. To save power, all functionalities on the SN are switched off when it is in silent mode. The energy usage of the SN can be regarded as zero in this state. In this work, the SNs are considered to be appropriate and the energy in severe conditions is restricted and divided. As an advantage, the SN will be feasible to fulfill the goal of maximizing life span.

Therefore, to save power usage, maximize the lifespan of the network, and ensure the performance of the network, the SN must reach a latent state as much as feasible when it is in the inactive monitoring phase.

### B. Energy Model

The energy consumption of the SN is influenced by a variety of variables. The primary elements that have a direct impact on the energy consumption of the SN are listed below. Different energy consumption aspects of SNs throughout the process of information need to be investigated to compute the power leftovers of the SN.

### C. Locust Search II

The LS-II approach is a metaheuristic method developed by studying the social behavior of locust swarms. The methods in LS-II prevent individuals from congregating in favorable locations and promoting the search area by redistributing the agents [25]. The LS-II technique combines both the tendencies to produce highly efficient domestic and global search capabilities. These properties have led to its application in solving difficult problems [26–43].

LS-II is a follow-up of the initial LS. Unlike the original LS, LS-II includes additional controllers and processes to boost its performance and reduce the accumulation of agents. Whenever the procedure arrives with multiple locally optimal solutions, such methods allow for a more accurate balance between exploratory and exploitative in identifying the global solution.

A population (PO) of N locusts constitutes a group of N possible solutions in LS-II, that is, PO ={ x1,..., xN}. The parts of PO communicate with one another as they explore an n-dimensional area. Every solution xi = [xi,1,..., xi,n] is located within the limited space UN = {xi Rn|ld xi,d ud (where ld and ud represent the minimum and maximum values for the d-th choice variables, respectively). LS-II, like any other metaheuristic technique, is based on an iterative development wherein the optimal solutions change their position with every repetition. The position of each candidate solution is changed by using several methods after the two behavioral patterns seen in the locust insects, namely, isolated and social activities.

### D. Solitary Process

Candidate solutions in the solo phase follow different paths in evaluating potential sources of food (good solutions). The search agents avoid focusing on the other elements during this procedure. This process is based on the concept of attractions and repulsion forces that are experienced by the answers in the PO populations. The ensuring attractiveness and repulsive pressures (called social force) are experienced by a given search agent $x_i$ under such parameters, which are explicitly described by the following formulation at each iteration t.

$$SP_i^t = \sum_{\substack{j=1 \\ i \neq j}}^{N} sp_{i,j}^t \tag{1}$$

Here $sp_{i,j}^t$ denotes the paired attraction-repulsion among the candidate answers $x_i$ and $x_j$, which are expressed by the representations.

$$sp_{i,j}^t = \rho(x_i, x_j) M(r_{i,j}) h_{i,j} + rand(-1,1) \tag{2}$$

opens $\rho(x_i, x_j)$ signifies the dominant level among $x_i$ and $x_j$. The fact is represented by the capable value as $M(r_{i,j})$, whereas $r_{i,j}$ represents the Euclidian distance $\|x_i - x_j\|$. The unit vector from $x_i$ to $x_j$ is the vector.

$$h_{i,j} = \frac{x_i - x_j}{\|x_i - x_j\|} \tag{3}$$

The below connection defines the social component $M(r_{i,j})$:

$$M(r_{i,j}) = Ae^{-r_{ij}/B} - e^{-r_{ij}} \tag{4}$$

Components A and B correspondingly denote the attraction-repulsion proportion and the influencing quantity. $\rho(x_i, x_j)$ is the function that represents the relative dominance of the candidate solutions of every search agent. It is evaluated with such a value inside the range [0, N-1] to execute $\rho(x_i, x_j)$. The strongest candidate option has the rank of 0, whereas the poorest component has the ranking of N-1. Concerning the generated fitness value, the ideas of the best and the worst are explored. The score of dominance is modeled as follows once the rankings for every candidate solution have been delegated:

$$\rho(x_i, x_j) = \begin{cases} e^{-\left(\frac{rank(x_i)}{N}\right)} if\ rank\ (x_i) \leq rank(x_j) \\ -e^{\left(\frac{rank(x_j)}{N}\right)} if\ rank\ (x_i) > rank(x_j) \end{cases} \tag{5}$$

As a result of the huge social power $S^t_i$ affects every search agent $x_i$ which has a distinct inclination to be attracted or rejected by the other components within the population's SP. In such cases, the new position $x_i^*$ estimated by the searching agent $x_i$ is due to the effect of the total force that calculated as follows:

$$x_i^* = x_i + S_i^t \tag{6}$$

*E. Social Process*

Even during the solitary phase, the social process is used to enhance the precision of the best candidate solution provided by $SP^*$. Even during the social practice, a subgroup of searching agents B = $\{b_1,..., b_q\}$ is produced, which includes the finest q parts of the population, $SP^*$. Next, within a constrained subspace $C_{iof} \in U$, a set Mi of new random solutions ($Mi = \{m^i_1,..., m^i_a\}$) is produced for every candidate solution $x^*_i \in B$. The boundaries of the subdomain $C_i$ wherein the alternative solutions are formed around $x^*_i = [\, x^i_1,..., x^i_{,n}\,]$ are shown as follows:

$$C^{lower}_{i,d} = x^*_{i,d} - r \quad (7)$$

$$C^{Upper}_{i,d} = x^*_{i,d} - r \quad (8)$$

The upper and the lower boundaries of every subcategory $C_i$ for the d decision variable ($d \in [1,..., n]$) are $C^{lower}_{i,d}$ and $C^{Upper}_{i,d}$, respectively. The perturbation r on the other side is computed using the standard formula,

$$r = \frac{\sum^n_{d=1}(u_d - l_d)}{n}.\beta \quad (9)$$

The lower and the upper bounds for the d- choice variables are denoted by $l_d$ and $u_d$, respectively. $n$is the total number of dimensions, and it correlates to a scaling component that controls the maximum size of $C_i$. The number of falls between the $beta\ ranges$is [0,1].

Eventually, the best component of the set incorporated by the value of $x^*_i$ with all its corresponding random solutions $\{m^i_1,..., m^i_a\}$ is acquired as the new role of the searching agent $x^{t+1}_i$. The following terminology can be used to describe this assignment.

$$x^{t+1}_i = best(x^*_i, m^i_1, \dots., m^i_a) \quad (10)$$

*F. Multi Swarm Optimization with Game Theory*

The multi-PSO swarms discussed in this paper do not generalize instantaneously because the swarms do not communicate. If the time is appropriate to acquire another swarm, there may be communication; nevertheless, one swarm diminishes as a result of a variety of information exchange topologies. Therefore, two multi-swarm approaches are proposed. Multi-charged particle swarm optimization (CPSO) is a multi-population form of CPSO. As a substitute for standard atom targets, multi-quantum swarm optimization uses quantum swarms based on quantization.

Blackwell founded CPSO on the orbit concept of an atom's nucleus, which orbits the earth by electrons. In CPSO, a certain number of atoms receive a charge that repels the other charged particles. The charged particles are those that have been assigned a value, while neutral particles are those that have not been given a charge. Particle speeds are updated using a standard updating formula with an additional option dependent on their closeness to the other electrons.

$$parent\overrightarrow{ve}_i(t+1) = \overrightarrow{\omega v}_i(t) + c_1\vec{r}_1(t)(p\overrightarrow{best} - \vec{x}_i(t) + c_2\vec{r}_2(t)(g\overrightarrow{best} - \vec{x}_i(t)) \quad (11)$$

Informula (11), $ve_i$ is the speed of the particle, $x_i$ is the location of the particle, $p_{best}$ is the individual perfect location of the particle, and $g_{best}$ is the best-found position of the swarm given a star neighborhood architecture.

The CPSO speed updating formula is as follows:

$$parent\overrightarrow{ve}_i(t+1) = \overrightarrow{\omega v}_i(t) + c_1\vec{r}_1(t)(p\overrightarrow{best} - \vec{x}_i(t) + c_2\vec{r}_2(t)\left(g\overrightarrow{best} - \vec{x}_i(t)\right) + \sum_{\forall j \neq i} a_{ij} \quad (12)$$

Eq. (13) gives the velocity among the particles i and j:

$$a_{ij} = \begin{cases} \frac{Qp_iQp_j}{\left\|\vec{x}_i - \vec{x}_j\right\|}\left(\vec{x}_i - \vec{x}_j\right) if\ \rho_{core} < \left\|\vec{x}_i - \vec{x}_j\right\| \leq p \\ 0 \quad otherwise \end{cases} \quad (13)$$

Here, $Qp_i$arethe charges of particle where i, p, and $p_{core}$ are the radii of influence of the accelerating component.

In vanilla PSO, neutral particles behave because they will. However, CPSO will show a convergence of neutral particles around $g_{best}$. The accelerating period following the divergence, wherein the charge carriers are rejected, causes more exploration and addresses the diversity loss problem [44–45]. As a result, CPSO provides a greater amount of variance throughout the loop and overcomes the linear collapsing issue that vanilla PSO has. Despite CPSO's excellent management of the diversity problem, an outside method to deal with the obsolete storage is required.

To explore the multiple potential peaks in parallel, a swarm is partitioned into sub-swarms. In concluding the research on the local optimum, diversity is raised but the probability is diminished. Furthermore, such swarms consist of two sorts of particles; i) PSO particles closely approximate the PSO method but aim to obtain a greater location and(ii) quantum particles orbit all around the sub swarm location in a radius cloud has to maintain diversity in addition to the optimization technique. The issue of diversity loss is addressed by quantum particles. The location of quantum particles is calculated using Eq. (14):

$$\vec{p}_i \in B_n(r_{cloud}) \quad (14)$$

The goal of MQSQ is to find a peak in every swarm, which is then monitored by the proposed method. A swarm interaction is performed in an eliminating process to ensure that the two swarms do not use the same peaks. The exclusion process among the swarms that are close together (distance smaller) employs a simple competitive strategy. The swarm with the best fitness value for its swarm attractor is declared as the winner, whereas for the loser, the swarming is dispatched and re-initialized in the search area.

*Gaming Model Establishment*

SNs can move between awake and asleep states reasonably to extend the lifespan of WSN. Consequently, SN dynamic switching can be considered as a game issue. The following is an example of the concept:

$$GT = \{, K, \{u_i\}\} \quad (15)$$

The players in WSN are represented by NP. The reason is that all the sensor network nodes are active in receiving and delivering the sensor data. All sensor network nodes are included in the game participants. To put it in another way, the game player is sensor $S_i$, with i = {1, 2, and n}.

The operational area of players is denoted by the letter K. In this paper, the sensor network evaluates the strategic space of the game player by evaluating the power usage. Specifically, the sensor network needs to enter the relaxed state from the active state but does not join the sleep state from the active state due to the sensible choice of SNs. The functional form of a player is represented by the UI.

$$u_i(s_i, s_{-i}) = U_i(s_i, s_{-i}) - CO_i(s_i, s_{-i}) \quad (16)$$

Here $s_i$ denotes the approach of the SN. Apart from the SNs, $s_i$ is the approach used by the nodes. The income value of SN $S_i$is $U_i$ ($s_i$,- $s_i$), whereas the cost value of $SNS_i$ is $CO_i$ ($s_i$, -$s_i$).

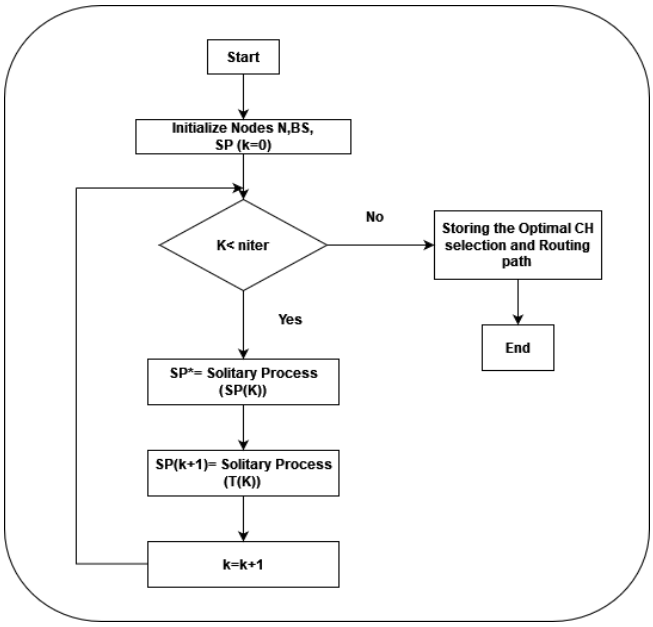



Figure. 2 Flowchart of MSGE-LS

In Fig.2 flowchart of the proposed method is shown for optimal CH selection and routing path. In the starting stage, it initializes the nodes, BS location and energy. Clusters are formed by calculating the distance between the networks that match the ground station and the energy levels of the nodes. MSO should be used to confirm the best local position. The LS-II completes the search method (ongoing process) that begins with the population's SP(k), that is, being randomly initialized during the first repetition (k = 0). The solo operator is then applied to the present population SP (k). As an outcome, an SP population (temporary population) is created. Lastly, the new population SP(k + 1) is generated using social operations. This cycle is repeated several times until it reaches a point.

## IV. Analysis and Results

The suggested MSGE-LS method is tested in this part by evaluating its efficiency with that of GA, LS, and MSO-based clustering. The statistical tests are run on a Dell OptiplexTM 3020 PC featuring an Intel Core i7-4770 3.4 GHz processor, 8 MB cache, inbuilt Intel graphics, 16 GB RAM, 1 TB HDD, and Windows 10 operating system. For the simulations, Matlab R2019a is employed. Tab.1 lists the network parameters that were employed during the process.

Table 1: Simulation parameters

| Parameters | Representations |
|---|---|
| Ranges between transmission | 30 m |
| Sensing distances among nodes | 10 m |
| Initial energy of nodes | 50 J |
| Bandwidth of nodes | 50 Kbps |
| Network's size | 400 sqm |
| Transmission rate of packets | 30 packets/s |
| Time taken for simulation | 32 min |
| Packet's size | 8,16,32,64,128,256 bytes |
| Mobility | 0.5–3.5 m/s |

The experiment of the proposed method is calculated using 200–1400 nodes and 0–800 rounds. The criteria that were taken into account are stated; number of clusters that generated the overall end-to-end delay in seconds, mean packet loss rates, lifespan computations, throughput, and total power dissipated in joules. The experimental calculations are shown in Tab. 2.

Table 2: Total number of clusters generated in MSGE-LS

| Total number of nodes | GA | LS | MSO | MSGE-LS |
|---|---|---|---|---|
| 200 | 10 | 11 | 13 | 15 |
| 400 | 15 | 16 | 18 | 21 |
| 600 | 24 | 23 | 25 | 29 |
| 800 | 27 | 28 | 27 | 32 |
| 1000 | 29 | 30 | 31 | 36 |
| 1200 | 29 | 31 | 33 | 40 |
| 1400 | 32 | 36 | 39 | 43 |

The proposed MSGE-LS method forms 4% of clusters higher than the other methods in 1400 rounds. The existing GA method forms 32 clusters, LS forms 36 clusters, and MSO forms 39 clusters in 1400 rounds. For 200, 400, 600, 800, 1000, 1200, and 1400 nodes, MSGE-LS improves approximately by 32%, 15%, 10%, 9%, 16%, and 13%, respectively, while compared with GA. MSGE-LS also outperforms LS by 9%, 14, 7%, 5%, 6% and 0. Fig.3 shows the number of clusters formed during 1400 nodes.

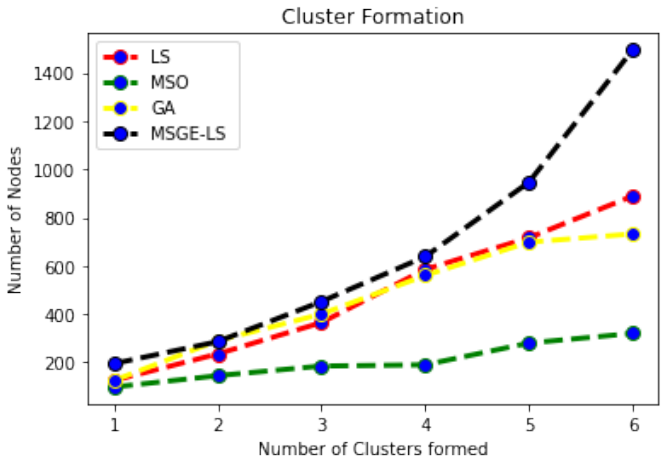



Figure. 3 Formations of clusters

The proposed MSGE-LS method achieves 65.31% of the minimum packet loss rate compared to the existing methods. In the existing methods, GA achieves 84.56%, LS achieves 80.35%, and MSO achieves 88.45% in average packet loss rates for 1400 rounds. For 200, 400, 600, 800, 1000, 1200, and 1400 nodes, the MSGE-LS has a reduced loss rate of 20%, 22%, 24%, 19%, 8%, and 20% when compared with GA and 12%, 11%, 18%, 17%, 11%, 9%, and 19% when compared to LS. MSGE-LS is also found to be more effective than MSO-based clustering by 6%, 5%, 5%, 7%, 4%, and 3% for 200, 400, 600, 800, 1000, and 1200 nodes, correspondingly. The experimental calculations are shown in Tab. 3. Fig.4 shows the average packet loss rate.

Table 3: Average packet loss rate

| Number of nodes used | GA | LS | MSO | MSGE-LS |
|---|---|---|---|---|
| 200 | 24.06 | 29.04 | 17.95 | 15.12 |
| 400 | 42.51 | 32.6 | 31.63 | 28.15 |
| 600 | 52.06 | 48.19 | 52.18 | 33.84 |
| 800 | 69.51 | 62.09 | 57.64 | 55.41 |
| 1000 | 73.56 | 71.42 | 64.54 | 51.45 |
| 1200 | 76.11 | 79.66 | 75.15 | 63.16 |
| 1400 | 84.56 | 80.35 | 88.45 | 65.31 |

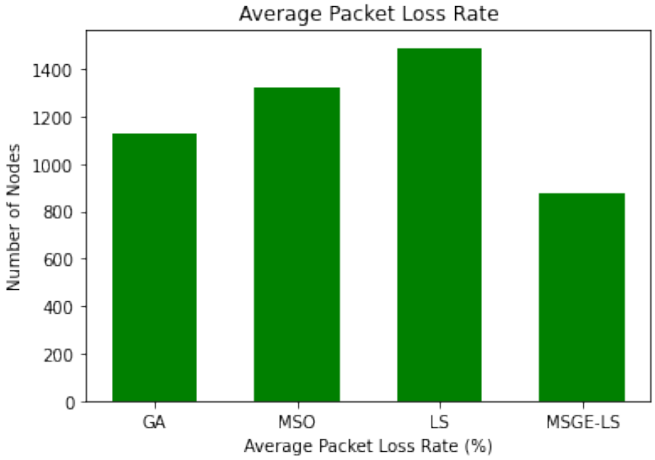



Figure. 4 Average packet loss rate

The proposed method has a minimum average end-to-end delay compared with GA, MSO, and LS. MSGE-LS achieve 0.057001s of end-to-end delay. Fig.5 shows the average end-to-end delay rates. When compared with GA, MSGE-LS has a reduced average of end-to-end delay benefit of 19%, 20%, 11%, 9%, and 5% for 200, 400, 600, 1000, and 1200 nodes, respectively, but a greater delay benefit of 9% for 800 nodes. When compared to MSO, MSGE-LS has a median end-to-end delay benefit of 0, 3%, 7%, 4%, 0, and 0.

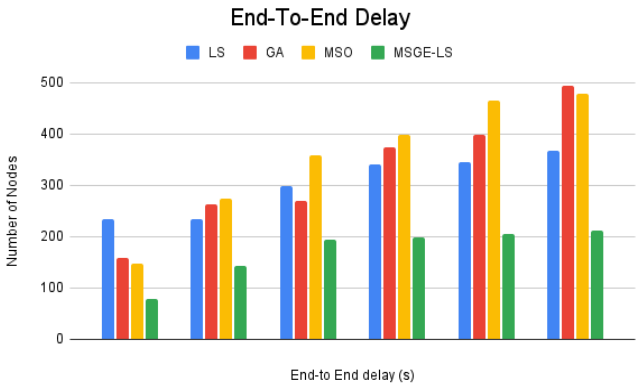



Figure. 5 End-to-end delay

The number of packets sent out to the ground station is measured by the throughput. It is preferable to transmit the greatest number of packages possible all through the lifespan of the network. Fig. 6 shows how to take this metric. The throughput is proportional to the number of active SNs sending the packets of data in each round, implying that the amount of data collected is greater. The suggested protocol provides the highest throughput, as shown in Fig. 6.

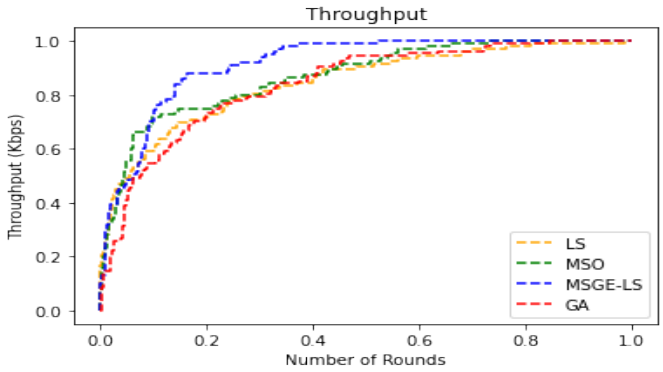



Figure. 6 Throughput

When compared to the existing techniques such as GA, LS, and MSO, MSGE-LS has a higher lifetime computation Fig.7. The lifetime computation of the proposed method is provided in Tab. 4.

Table 4: Lifetime computation

| Number of nodes | Number of Rounds | LS | GA | PSO | MSGE-LS |
|---|---|---|---|---|---|
| 600 | 100 | 945 | 955 | 840 | 1031 |
| 800 | 200 | 998 | 992 | 978 | 1098 |
| 1000 | 300 | 1050 | 1170 | 975 | 1297 |
| 1200 | 400 | 1120 | 1195 | 1060 | 1335 |
| 1400 | 500 | 1185 | 1235 | 1190 | 1496 |

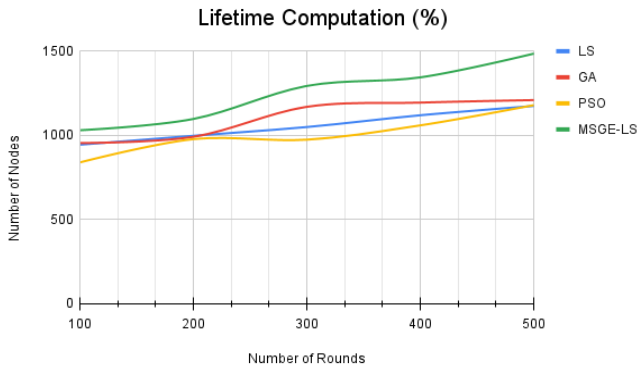



Figure. 7 Lifetime computation

The cost of computation is directly proportional to time. If the minimum nodes take more time to cluster, they will consume more energy and have a high storage cost. In this proposed research, the computation time is of high concern, and the cost of computation is reduced. This helps in improving the lifetime of the network with less component loss in terms of battery life, storage space and so on.

## V. Conclusion

Two of the most essential criteria for operating WSNs successfully and consistently are energy conservation and energy consumption. At present, more importance is given to technical issues and resource management. The main goal of this research is to use the LS-II method to improve WSN CH selection. Therefore, a method that relies on MSO and game theory has been developed to improve WSN and CH selection routing. The solution generated was based on both energy efficiency and network responsiveness. The suggested method optimizes routing and extends the lifetime of the network by selecting the effective CHs in reaction to the networks. According to the findings, the suggested scheme outperforms the comparative methods in simulations. The suggested protocol is evaluated using several parameters, such as the lifetime of the network, energy usage, energy divergence, number of packets transmitted, and system stability duration. The research has shown that the suggested methodology performs well in terms of increasing the lifespan of the network and providing a good energy load.

The limitation of this research is the lack of energy efficiency across the network. In such a situation, the challenging due nodes are considered as players. If one node acts as a player, the whole algorithm takes responsibility only for that player's node. In the future, a better optimization model can be used for node optimization.

## Acknowledgment

This research was supported by Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (P0012724, The Competency Development Program for Industry Specialist) and the Soonchunhyang University Research Fund.

## References

[1] S.Suganthi, N.Umapathi, M.Mahdal and M.Ramachandran, "Multi swarm optimization based clustering with tabu search in wireless sensor network," *Sensors*, vol.22, no.5, pp.1736-1745,2022.

[2] A.Rodríguez, M.P.Cisneros, J.C.R.Caro, C.D.V.Soto, J.Gálvez *et al*., "Robust clustering routing method for wireless sensor networks considering the Locust search scheme," *Energies*, vol.14,no.11, pp.3019-3025,2021.

[3] X.Yan, C.Huang, J.Gan and X.Wu, "Game theory based energy-efficient clustering algorithm for wireless sensor networks," *Sensors*, vol.22, no.2, pp.478-485,2022.

[4] X.Huan, K.S.Kim, S.Lee, E.G.Lim and A. Marshall, "A beaconless asymmetric energy-efficient time synchronization scheme for resource-constrained multi-hop wireless sensor networks," *IEEE Transactions on Communications*, vol.68, no.3, pp.1716-1730, 2019.

[5] S.Chamanian, S.Baghaee, H.Uluşan, O.Zorlu, E.U. Biyikoglu *et al*., "Implementation of energy-neutral operation on vibration energy harvesting WSN, "*IEEE Sensors Journal*, vol.19, no.8, pp.3092-3099, 2019.

[6] R.Ashween, B.Ramakrishnanand and M.M. Joe, "Energy efficient data gathering technique based on optimal mobile sink node selection for improved network lifetime in wireless sensor network (WSN),"*Wireless Personal Communications*, vol.113, no.4, pp.2107-2126, 2020.

[7] S.Q.Zhang, Y.Tao and J.J.Dai, "Multi-hop clustering routing protocol for energy harvesting wireless sensor networks," *Computer Engineering Design*, vol.40, no. 4, pp.611-616, 2019.

[8] N.Saxena, A.Roy and J.Shin, "Dynamic duty cycle and adaptive contention window based QoS-MAC protocol for wireless multimedia sensor networks," *Computer Networks*, vol.52, no.13, pp.2532-2542, 2018.

[9] M.B.Dowlatshahi, M.K.Rafsanjani and B.B.Gupta, "An energy-aware grouping memetic algorithm to schedule the sensing activity in WSNs-based IoT for smart cities," *Applied Soft Computing*, vol.108, no.2, pp.107473, 2021.

[10] S.Kassan, J.Gaber and P.Lorenz, "Game theory based distributed clustering approach to maximize wireless sensors network lifetime," *Journal of Network and Computer Applications*, vol.123, no. 11, pp.80-88, 2018.

[11] S.Shen, H.Ma, E.Fan, K.Hu, S.Yu *et al*., "A non-cooperative non-zero-sum game-based dependability assessment of heterogeneous WSNs with malware diffusion," *Journal of Network and Computer Applications*, vol.91, no.3, pp.26-35, 2017.

[12] A.González, E.Cuevas, F.Fausto, A.Valdivia and R.Rojas, "A template matching approach based on the behavior of swarms of locust," *Applied Intelligence*, vol.47, no.4, pp.1087-1098, 2017.

[13] X.Fu, P.Pace, L.Aloi Yang and G.Fortino, "Topology optimization against cascading failures on wireless sensor networks using a memetic algorithm," *Computer Networks*, vol.177, no.3, pp.107327, 2020.

[14] X.Cheng, D.Ciuonzo and P.S.Rossi, "Multibit decentralized detection through fusing smart and dumb sensors based on Rao test," *IEEE Transactions on Aerospace and Electronic Systems*, vol.56,no.2, pp.1391-1405,2019.

[15] D.Ciuonzo, P.S.Rossi and P.K.Varshney, "Distributed detection in wireless sensor networks under multiplicative fading via generalized score tests," *IEEE Internet of Things*, vol.8,no.11, pp.9059-9071, 2021.

[16] A.Shahraki, A.Taherkordi, O.Haugen and F.Eliassen, "Clustering objectives in wireless sensor networks: A survey and research direction analysis," *Computer Networks*, vol.180, no.4, pp.107376, 2020.

[17] M.Krishnan, S.Yun and Y.M.Jung, "Improved clustering with firefly-optimization-based mobile data collector for wireless sensor networks," *AEU-International Journal of Electronics and Communications*, vol.97, no. 3, pp.242-251, 2018.

[18] M.Krishnan, Y.M.Jung and S.Yun, "An improved clustering with particle swarm optimization-based mobile sink for wireless sensor networks," in *Proc.2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, pp.11-12, 2018.

[19] Y.Xu, O.Ding, R.Qu and K.Li, "Hybrid multi-objective evolutionary algorithms based on decomposition for wireless sensor network coverage optimization," *Applied Soft Computing*, vol.68, no. 4, pp.268-282,2018.

[20] T.Kaur and D.Kumar, "Particle swarm optimization-based unequal and fault tolerant clustering protocol for wireless sensor networks," *IEEE Sensors Journal*, vol.18,no.11, pp.4614-4622,2018.

[21] V.Varsha, J.Singh and M.Bala, "Tabu search based energy efficient clustering protocol for wireless sensor networks," *Global Journal of Computers & Technology*, vol.5,no.2, pp.302-309,2017.

[22] G.P.Gupta, "Improved cuckoo search-based clustering protocol for wireless sensor networks," *Procedia Computer Science*, vol.125, no. 6, pp.234-240, 2018.

[23] L.Kong, J.S.Pan, V.Snášel, P.W.Tsai and T.W.Sung, "An energy-aware routing protocol for wireless sensor network based on genetic algorithm," *Telecommunication Systems*, vol.67, no.3, pp.451-463, 2018.

[24] J.Wang, Y.Gao, W.Liu, A.K.Sangaiah and H.J.Kim, “An improved routing schema with special clustering using PSO algorithm for heterogeneous wireless sensor network,” *Sensors*, vol.19,no.3, pp.671-684,2019.

[25] M. Abouhawwash and K. Deb, "Karush-Kuhn-tucker proximity measure for multi-objective optimization based on numerical gradients," in *Proc. 2016 on Genetic and Evolutionary Computation Conference Companion, ACM*, Denver Colorado USA, pp. 525-532, 2016.

[26] A.H. El.Bassiouny, M. Abouhawwash and H.S. Shahen, "New generalized extreme value distribution and its bivariate extension," *International Journal of Computer Applications*, vol. 173, no. 3, pp. 1-10, 2017.

[27] A.H.El.Bassiouny, M.Abouhawwash and H.S.Shahen, "Inverted exponentiated gamma and it’s a bivariate extension," *International Journal of Computer Application*, vol. 3, no. 8, pp. 13-39, 2018.

[28] A.H.El.Bassiouny, H.S.Shahen and M.Abouhawwash, "A new bivariate modified Weibull distribution and its extended distribution," *Journal of Statistics Applications & Probability*, vol. 7, no.2, pp. 217-231, 2018.

[29] M. Abouhawwash and M.A. Jameel, "KKT proximity measure versus augmented achievement scalarization function," *International Journal of Computer Applications*, vol. 182, no. 24, pp. 1-7, 2018.

[30] H.S.Shahen, A.H.El.Bassiouny and M.Abouhawwash, "Bivariate exponentiated modified Weibull distribution," *Journal of Statistics Applications & Probability*, vol. 8, no. 1, pp. 27-39, 2019.

[31] M. Abouhawwash and M.A.Jameel, "Evolutionary multi-objective optimization using Benson’s kush-Kuhn-tucker proximity measure," in *Proc.International Conference on Evolutionary Multi-Criterion Optimization*, East Lansing, Michigan, USA, pp. 27-38, 2019.

[32] M. Abouhawwash, M.A. Jameel and K. Deb, "A smooth proximity measure for optimality in multi-objective optimization using Benson’s method," *Computers& Operations Research*, vol. 117, no. 2, pp. 104900, 2020.

[33] M.Abouhawwash, K.Deb and A. Alessio, "Exploration of multi-objective optimization with genetic algorithms for PET image reconstruction," *Journal of Nuclear Medicine*, vol. 61, no. 1, pp. 572-572, 2020.

[34] M.AbdelBasset, R. Mohamed, M. Elhoseny, M. Abouhawwash, Y. Nam *et al*., "Efficient MCDM model for evaluating the performance of commercial banks: A case study," *Computers, Materials & Continua*, vol. 67, no. 3, pp. 2729-2746, 2021.

[35] B.Gomathi, S.Balaji, V.K.Kumar, M.Abouhawwash, S. Aljahdali *et al*., "Multi-objective optimization of energy aware virtual machine placement in the cloud data center," *Intelligent Automation & Soft Computing*, vol.33, no.3, pp. 1771-1785, 2022.

[36] M.Kumar, K.Venkatachalam, M. Masud and M. Abouhawwash, "Novel dynamic scaling algorithm for energy-efficient cloud computing," *Intelligent Automation & Soft Computing*, vol.33, no.3, pp. 1547-1559, 2022.

[37] R.S.Ram, K.Venkatachalam, M.Masud and M.Abouhawwash, "Air pollution prediction using dual graph convolution LSTM technique," *Intelligent Automation & Soft Computing*, vol.33, no.3, pp. 1639-1652, 2022.

[38] A.J. Basha, N. Rajkumar, M.A. AlZain, M. Masud and M. Abouhawwash, "Fog based self-sovereign identity with RSA in securing IoMT data," *Intelligent Automation & Soft Computing*, vol. 34, no. 3, pp.1693-1706, 2022.

[39] G. Ravikumar, K. Venkatachalam, M.A. AlZain, M. Masud and M. Abouhawwash, "Neural cryptography with fog computing network for health monitoring using IoMT," *Computer Systems Science and Engineering*, vol. 44, no.1, pp.945-959, 2023.

[40] R. Rajdevi, K. Venkatachalam, M. Masud, M.A. AlZain and M. Abouhawwash, "Proof of activity protocol for IoMT data security," *Computer Systems Science and Engineering*, vol. 44, no. 1, pp.339-350, 2023.

[41] N. Mittal, H. Singh, V. Mittal, S. Mahajan, A.K. Pandit *et al*., "Optimization of cognitive radio system using self-learning salp swarm algorithm," *Computers, Materials & Continua*, vol.70, no.2, pp.3821-3835, 2022.

[42] A.Garg, A.Parashar, D.Barman, S.Jain, D.Singhal *et al*., "Autism spectrum disorder prediction by an explainable deep learning approach," *Computers, Materials & Continua*, vol.71, no.1, pp.1459-1471, 2022.

# Feature Fusion Based Deep Transfer Learning Based Human Gait Classification Model

C.S.S. Anupama[1], Rafina Zakieva[2], Afanasiy Sergin[3], E Laxmi Lydia[4], Seifedine Kadry[5,6,7], Chomyong Kim[8] and Yunyoung Nam[8, *]

[1]*Department of Electronics and Instrumentation Engineering, V.R.Siddhartha Engineering College, Vijayawada 520007, India*
[2]*Candidate of Pedagogical Sciences, Department of Industrial electronics and lighting engineering, Kazan State Power Engineering University, Kazan, 420066, Russia*
[3]*Department of Theories and Principles of Physical Education and Life Safety, North-Eastern Federal University named after M.K. Ammosov, Yakutsk, 677000, Russia*
[4]*Department of Computer Science and Engineering, GMR Institute of Technology, Andhra Pradesh, Rajam, India*
[5]*Department of Applied Data Science, Noroff University College, Kristiansand, Norway*
[6]*Artificial Intelligence Research Center (AIRC), Ajman University, Ajman, 346, United Arab Emirates*
[7]*Department of Electrical and Computer Engineering, Lebanese American University, Byblos, Lebanon*
[8]*Department of ICT Convergence, Soonchunhyang University, Asan, 31538, Korea*
[*]*Corresponding Author: Yunyoung Nam. Email: ynam@sch.ac.kr*

*Abstract*—**Gait is a biological typical that defines the method by that people walk. Walking is the most significant performance which keeps our day-to-day life and physical condition. Surface electromyography (sEMG) is a weak bioelectric signal that portrays the functional state between the human muscles and nervous system to any extent. Gait classifiers dependent upon sEMG signals are extremely utilized in analysing muscle diseases and as a guide path for recovery treatment. Several approaches are established in the works for gait recognition utilizing conventional and deep learning (DL) approaches. This study designs an Enhanced Artificial Algae Algorithm with Hybrid Deep Learning based Human Gait Classification (EAAA-HDLGR) technique on sEMG signals. The EAAA-HDLGR technique extracts the time domain (TD) and frequency domain (FD) features from the sEMG signals and is fused. In addition, the EAAA-HDLGR technique exploits the hybrid deep learning (HDL) model for gait recognition. At last, an EAAA-based hyperparameter optimizer is applied for the HDL model, which is mainly derived from the quasi-oppositional based learning (QOBL) concept. A brief classifier outcome of the EAAA-HDLGR technique is examined under diverse aspects, and the results indicate improving the EAAA-HDLGR technique. The results imply that the EAAA-HDLGR technique accomplishes improved results with the inclusion of EAAA on gait recognition.**

## I. INTRODUCTION

Human gait, how an individual walks, is personally distinctive because of its physical properties difference betwixt individuals and might be employed as a novel biometric for the authentication and identification of a person [1]. In comparison to other biometrics, namely iris, face, and fingerprint, human gait recognition (HGR) has the advantage of non-invasion, non-cooperation (without any cooperation or interaction with the subject), hard to disguise and long distance, making it very attractive as a means of detection and demonstrates huge potential in the applications of surveillance and security [2, 3]. Many sensing modalities involving wearable devices, vision, and foot pressure were used for capturing gait data [4].

Traditional HGR techniques include data preprocessing and features extracted in a handcrafted manner for additional identification [5], frequently suffer from various challenges and constraints enforced by the difficulty of the tasks, namely occlusions, viewing angle, locating the body segments, shadows, large intra-class variations, etc. [6, 7]. Emerging trends in machine learning (ML), called deep learning (DL), have become apparent in the past few years as a ground-breaking tool for handling topics in computer vision, speech, sound, and image processing, tremendously outperforming virtually any baseline established previously [8]. The new model exempts the requirement of manually extracting representative features from the expert and provides primary outcomes based on HGR, surpassing the present difficulties and opening room for additional investigation [9, 10]. The manually extracted feature was affected by the smartphone's position while gathering the information [11]. Consequently, some typical statistical manually extracted feature was collected from raw smartphone sensor information. Afterwards, the extraction of handcrafted features, the shallow ML classifier, was used to identify many physical activities of the human. Therefore, shallow ML algorithm relies on handcrafted feature [12, 13]. The DL algorithm is more advanced than the shallow ML algorithm since the DL algorithm automatically learns useful features from the raw sensory information without human intervention and identifies the human's physical activities [14]. The shallow ML algorithm with DL algorithms and handcrafted extracting features with automatically learned features achieved better outcomes in carrying out the smartphone-based HGR model. So, it is apparent that integrating manually extracted features with automatically learned features in the DL algorithm might enhance the potential of the smartphone-based HGR paradigm [15].

Khan et al. [16] developed a lightweight DL algorithm for the HGR method. The presented algorithm involves sequential step–pretrained deep model selection of feature classification. Initially, two lightweight pretrained methods

are considered and finetuned with respect to other layers and freeze some middle layers. Thereafter, the model was trained by means of deep transfer learning (DTL) algorithm, and the feature was engineered on average pooling and fully connected layers. The fusion is carried out by means of discriminative correlation analysis, enhanced by the moth-flame optimization (MFO) technique. Liang et al. [17] examine the effect of every layer on parallel infrastructured CNN. In particular, we slowly freeze the parameter of GaitSet from high to low layers and see the performance of finetuning. Furthermore, the rise of the frozen layer has negative consequences on the performance; it could reach the maximal efficacy with a single convolution layer unfrozen.

The author in [18] proposed a novel architecture for HGR using DL and better feature selection. During the augmentation phase, three flip operations were performed. During feature extraction phase, two pre-trained models have taken place, NASNet Mobile and Inception-ResNetV2. These two models were trained and fine-tuned by means of the TL algorithm on the CASIA B gait data. The feature of the selected deep model was enhanced by means of an adapted three-step whale-optimized algorithm, and the better feature was selected. Hashem et al. [19] proposed an accurate and advanced end-user software system that is capable of identifying individuals in video based on the gait signature for the purpose of hospital security. TL algorithm based on pretrained CNN was used and capable of extracting deep feature vectors and categorizing people directly rather than a typical representation that includes hand-crafted feature engineering and computing the binary silhouettes.

In [20], proposed a novel fully automatic technique for HGR under different view angles using the DL technique. Four major phases are included: recognition using supervised learning methods, pre-processing of the original video frame, which exploits pre-trained Densenet-201 CNN model for extracting features, and decrease of further features in extraction vector based on a hybrid selection technique. Sharif et al. [21] suggest a method that efficiently deals with the problem related to walking styles and viewing angle shifts in real time. The subsequent steps are included: (a) feature selection based on the kurtosis-controlled entropy (KcE) method, then a correlation-based feature fusion phase, (b) realtime video capture, and (c) extraction feature utilizing TL on the ResNet101 deep model. Then, the most discriminative feature was categorized by means of the innovative ML classifier.

This study designs an Enhanced Artificial Algae Algorithm with Hybrid Deep Learning based Human Gait Classification (EAAA-HDLGR) technique on sEMG signals. The EAAA-HDLGR technique initially extracts the time domain (TD), and frequency domain (FD) features from the sEMG signals and is fused together. In addition, the EAAA-HDLGR technique exploits the hybrid deep learning (HDL) model for gait recognition. At last, an EAAA-based hyperparameter optimizer is applied for the HDL model, smainly derived from the quasi-oppositional based learning (QOBL) concept. A brief classifier outcome of the EAAA-HDLGR technique is examined under diverse aspects.

## II. The Proposed Model

In this study, we have derived a new EAAA-HDLGR technique for gait recognition using sEMG signals. Primarily, the EAAA-HDLGR technique derived the TD as well as FD features from the sEMG signals, which are then fused together. In addition, the EAAA-HDLGR technique exploited the HDL model for gait recognition. At last, an EAAA-based hyperparameter optimizer is applied for the HDL model, mainly derived from the QOBL concept. Fig. 1 depicts the workflow of the EAAA-HDLGR algorithm.

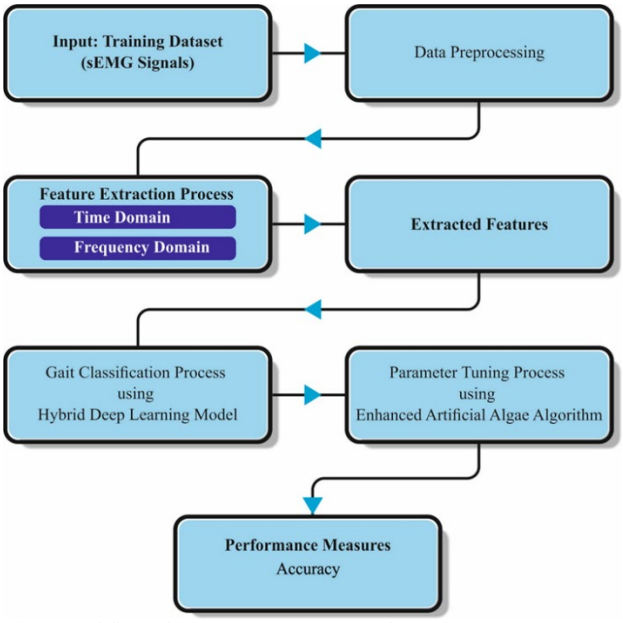



Fig. 1 Workflow of EAAA-HDLGR approach

### A. Feature Extraction Process

Afterwards, de-noising, the TD and FD features of all the channels of the EMG signal can be extracted [22]. During this case, the 3 representative time domain features comprising variance (VAR), zero crossing points (ZC), and mean absolute value (MAV) can be utilized as frequency domain features. MAV gets the benefit of properties in which sEMG signals are huge amplitude fluctuations from the time domain that are linearly compared with muscle activation level. The maximum value of $MAV$ is the superior activation level of muscles.

$$MAV = \frac{1}{N}\sum_{k=1}^{N} |x_k| \quad (1)$$

Whereas, $x_k (k = 1,2, \dots, N)$ implies the sEMG time series with a window length of $N$. VAR is the size of signal powers of sEMG signals and is formulated as:

$$VAR = \frac{1}{N-1}\sum_{k=1}^{N} x_k^2 \quad (2)$$

$ZC$ implies the count of times the sEMG waveform passes with the zero point to avoid signal cross-calculating due to low-level noise. It can be a mathematical process as follows:

$$ZC = \sum_{k=1}^{N} sgn(-x_k x_{k+1}) \quad (3)$$

Whereas, $sgn(x) = \begin{cases} 1\ x > 0 \\ 0\ otherwise \end{cases}$

It can choose 2 representative frequency domain features like average power frequency $f_{mean}$ and median frequency $f_{mf}$ determined as:

$$f_{mean} = \frac{\int_0^{+\infty} fP(f)df}{\int_0^{+\infty} P(f)df} \quad (4)$$

$$\int_0^{f_{mf}} P(f)df = \int_{f_{mf}}^{+\infty} P(f)df = \frac{1}{2}\int_0^{+\infty} P(f)df \quad (5)$$

Whereas $P(f)$ implies the power spectral density of sEMG signals, and $f$ denotes the frequency.

### B. Gait Classification using HDL Model

During this study, the HDL technique was employed for the gait classifier. It comprises CNN with LSTM and Bi-LSTM technique [23]. CNN is increasingly popular under the domains of DL and is contained in convolution (Conv) and pooling layers. The Conv layer function is to remove useful features in the input database. Its interior includes one or more Conv kernels. The Conv layer was executed to remove effectual features by sliding the Conv kernel on the feature. Afterwards, enhance a Max-Pooling layer and then the Conv layer; the Max-Pooling layer maintains the strong features and extracts the weaker feature to prevent over-fitting and decrease the complexity. LSTM avoids gradient vanishing and explosion from the trained process of RNN, whereas input, output, and forget gates can be projected. The 3 gate functions effectively solve the trained problem of RNN. The input gate determines the data that existing time carries to future time. The forget gate identifies the data count of preceding times which is preserved in the current period. The resultant gate chooses the resultant of existing to future states. The succeeding equation illustrates the equation of distinct cells from LSTM:

$$i_z = \sigma(W_{xi}x_z + W_{hi}h_{z-1} + b_i) \quad (6)$$
$$f_z = \sigma\left(W_{xf}x_z + W_{hf}h_{z-1} + b_f\right) \quad (7)$$
$$o_z = \sigma(W_{xo}x_z + W_{ho}h_{z-1} + b_o) \quad (8)$$
$$\tilde{C}_z = \tanh\left(W_{xc}x_z + W_{hc}h_{z-1} + b_c\right) \quad (9)$$
$$c_z = f_z e c_{z-1} + i_z e \tilde{C}_z) \quad (10)$$
$$h_z = o_z e \tanh\left(c_z\right) \quad (11)$$

At this point, $W_{hf}, W_{hc}, W_{ho}, W_{hi}$ implies the weighted in hidden layer to forget, memory cell, output, and input gates. $W_{xf}, W_{xo}, W_{xi}$, and $W_{xc}$ signify the weight of forgetting gate, output gate, input gate, and memory cell. $f_z, i_z, 0_z$ illustrate the $z^{th}$ forget, input, and output gates. $b_c\ b_o, b_f, b_i$ stands for the bias values of a memory cell, output gate, forget gate, and input gate. Bi-LSTM is a group of forward and backward LSTM that could provide relevant data in forward and backward ways and concatenate the prediction. LSTM is suitable the time-related data in one way. BiLSTM improves the opposite way of LSTM in such a method that BiLSTM captures the pattern, which LSTM ignores. LSTM-BiLSTM classifier the gait based on extracting feature.

- CNN was projected with three 1-D Conv layers, and the count of Conv kernels is set as 16, 32, 64, correspondingly. The Conv kernel size is set as 2, and the striding phase has set as. The effectual feature was extracted with stride from the Conv kernel. Then, add the Max-Pooling layer for all the Conv layers; afterwards, the pooling window size was set as 2, and the stride stage was set as one. The Max-Pooling layer decreases the feature effort and avoids overfitting.
- By assigning weight to features using the attention process, the attention block enhances the outcome of time series features, restrains the interference of insignificant features, and effectually explains this problem that process isn't judge the effect of vital of various time series features.
- The extraction feature assumed that input of 1 Bi-LSTM layer and 2 LSTM layers achieved the classifier outcome.

### *C. Parameter Optimization using EAAA Model*

To improve the recognition rate, EAAA based hyperparameter optimizer is applied to the HDL model. Artificial algae are generally known to describe the features of algae and demonstrate that they can be responsive to solutions from the problem spaces [24]. In relation to a real one, artificial algae demonstrate that if it is implemented in environment by moving to light source for photosynthesizing by spiral swimming, it is able to switch superior species and eventually multiply with mitotic division. Thus, this procedure contains 3 important processes Helical Motion, Evolutionary Process, and Adaptation. The term algae colony signifies the collection of algae cells that lives together. Algae colony and population can demonstrate in the subsequent formula.

$$Population = \begin{bmatrix} \chi_1^1 & \cdots & \chi_1^D \\ \vdots & \ddots & \vdots \\ \chi_N^1 & \cdots & \chi_N^D \end{bmatrix} \quad (12)$$
$$Algae\ colony = [\chi_i^1, \chi_i^2\ \chi_i^D] \quad (13)$$

The algae colony work as separate cell and permits together, and the cell from the colony can pass away an unfavorable level of life. The colony present at the optimal point is named a better colony, and it includes better algae cells. The development kinetics of algae colonies calculated by Monod approach as demonstrated under.

$$\mu = \frac{\mu \max S}{K_s + S} \quad (14)$$

In Eq. (14), $\mu$ signifies the rate of growth, $\mu$ max implies the maximal rate of growth, $S$ represents the nutrient concentration, and $K$ is divided as to algal colony. In Eq. (17), $\mu$max value is set to one (dependent upon mass conservation model, an entire count converted as to biomass can correspond to the count of substrate expended to all the time units. The size of $i^{th}$ algae colonies at time $t+1$ from the Monod equation is depicted under:

$$G_i^{t+1} = \mu_i^t G_i^t\ i = 1,2,N \quad (15)$$

In Eq. (15), $G_i^t$ denotes the size of $i^{th}$ algae colonies at $t$ time, $N$ represents the amount of algae colonies. The algae colony offers optimum solution that develops higher as a sign of greater count of nutrients is obtained. In all the algae cells, the minimal algae colony dies from the evolution procedure, and the algae cell of maximum algae colony reproduces. Such procedure is complete in the subsequent formulas:

$$biggest^t = \max G_i^t i = 1,2,\cdots N \quad (16)$$
$$smallest^t = \min G_i^t i = 1,2,\cdots N \quad (17)$$
$$smallesl_m^t = biggest_m^t m = 1,2,\cdots D\ D \quad (18)$$

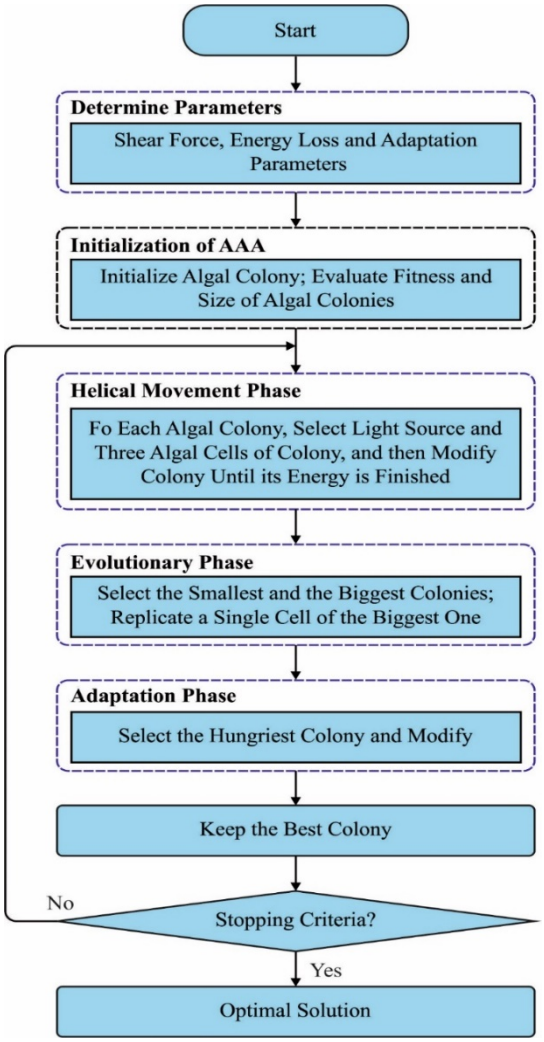


Fig. 2 Flowchart of AAA

In the above formula, $D$ signifies that there is a problem feature, whereas the maximum describes the largest colony of algae, and the minimal symbolizes the lower colony of algae. A primary hunger value is 0 for every artificial alga, and Adaptation process stops by an alteration in the hunger level.

$$starving^t = \max A_i^t i = 1,2,\cdots N \quad (19)$$

$$starving^{t+1} = slarving^t + (bigges^t - starving^t)rand \quad (20)$$

During this equation, the value $A$ signifies the hunger value of $i^{th}$ algae colonies at $t$ time, and $\text{starving}^t$ describes the algae colonies with maximal angle value at $t$ time. In Eq. (20), the variation parameter states the adaptation scheme and is utilized at $t$ time. Generally, the variation parameter value lies in the interval of 0 and 1. In AAA, the movement of algae cells are spiral, the viscous drag is realized as shear force commensurate to algae cell size, gravity restricting movement is presented as zero, and viscous drag demonstrates as zero. At this time, $\tau(X)$ implies the friction surface, the friction surface is the surface area, and Algae colony has spherical from the shape demonstrated in the under equation. Fig. 2 depicts the flowchart of AAA.

$$\tau(x_i) = 2\pi r^2 \quad (21)$$

$$\tau(x_i) = 2\pi\left(3\sqrt{\frac{3G_i}{4\pi}}\right)^2 \quad (22)$$

The distance to light source and friction surface is to control the step size of motions.

$$x_{im}^{t+1} = x_{im}^t + (x_{jm}^t - x_{im}^t)(\Delta - \tau^t(X))p \quad (23)$$

$$x_{ik}^{t+1} = x_{ik}^t + \left(x_{jk}^t - x_{ik}^t\right)\left(\Delta - \tau^t(x_i)\right)\cos\alpha \quad (24)$$

$$x_{i1}^{t+1} = x_{i1}^t + \left(x_{j1}^t - x_{i1}^t\right)\left(\Delta - \tau^t(x_i)\right)\sin\beta \quad (25)$$

For the helical rotation of algae cells, ($x_{im}^t$, $x_{ik}^t$, and $x_{i1}^t$) are stated as the co-ordinate of the algae cell ($x$, $y$, and z) at $t$ time. $\alpha$ and $\beta \in [0,2]$; $p \in [-1,1]$; $\Delta$ the force to procedure; $t$ (xi) signifies the friction surface area of $i^{th}$ algae cells. The AAO technique arbitrarily creates the primary solution in the search range. Once the solution is superior to the global better solution, afterward, the convergence population rate can be slower and simply decrease as to the local better solution. Tizhoosh projected oppositional-based learning (OBL) technique for preventing these local better problems [25].

Also, the QOAAO technique resulted by utilizing quasi-opposition based population initialization. At this point, the random solution was higher than global better solution if related to their opposite solution. So, the N better individual was chosen in the distinct population containing $N$ arbitrary individuals and the opposite solution of individuals. Let $X = (X_1, X_2, \ldots, X_D)$ refers to the point in $d$ dimension space, and its opposite point signifies $X^{OBL} = (X_1^{OBL}, X_2^{OBL}, \ldots, X_D^{OBL})$ which is calculated by the subsequent formula; and its quasi-opposite point $X^{QOBL} = \left(X_1^{QOBL}, X_2^{QOBL}, \ldots, X_D^{QOBL}\right)$ was measured as:

$$X_d^{OBL} = lb_d + ub_d - X_d, \quad (26)$$

$$X_d^{QOBL} = \begin{cases} \frac{lb_d+ub_d}{2} + rand(0,1) \times (X_d^{OBL} - \frac{lb_d+ub_d}{2}) X_d < \frac{lb_d+ub_d}{2} \\ X_d^{OBL} + rand(0,1) \times (\frac{lb_d+ub_d}{2} - X_d^{OBL}) X_d \geq \frac{lb_d+ub_d}{2} \end{cases}, \quad (27)$$

In the formula, the $d$ dimension vector of $X$ was signified as $X_d$, $X_d^{OBL}$, and $X_d^{QOBL}$ denotes the $d$ dimensional opposite point $X^{OBL}$, and quasi-opposite point $X^{QOBL}$ correspondingly. Also, $lb_d$ and $ub_d$ illustrates the $d$ dimensional of lower as well as upper limits of problems, correspondingly.

The fitness chosen is a vital feature in the EAAA system. The solution encoder was exploited to assess the aptitude (goodness) of candidate solutions. At this point, the accuracy value is main form employed to design a fitness function.

$$Fitness = \max(P) \quad (28)$$

$$P = \frac{TP}{TP+FP} \quad (29)$$

From the expression, TP represents the true positive, and FP denotes the false positive value.

## III. Results and Discussion

In this section, the gait classification results of the EAAA-HDLGR approach are investigated in detail. Tab. 1 and Fig. 3 demonstrate the overall gait classification outcomes of the EAAA-HDLGR model with other ML models on TD features [22]. The experimental results inferred that the EAAA-HDLGR model exhibits effectual outcomes. For instance, with pre-stance stage, the EAAA-HDLGR model has highlighted increasing $accu_y$ of 96.07% while the SVM, ELM, LSTM, DBN, and SSA-DBN models have attained reducing $accu_y$ of 95%, 96.07%, 93.49%, 93.63%, and 96.99% respectively. Meanwhile, with Terminal-Stance stage, the EAAA-HDLGR approach has emphasized increasing $accu_y$ of 98.95% while the SVM, ELM, LSTM, DBN, and SSA-DBN models have attained decreasing $accu_y$ of 92.01%, 95.50%, 97.33%, 94.73%, and 95.78% correspondingly. Finally, with Terminal-Swing stage, the

EAAA-HDLGR technique has demonstrated higher $accu_y$ of 98.03% while the SVM, ELM, LSTM, DBN, and SSA-DBN methods have attained lesser $accu_y$ of 96.92%, 94.79%, 92.76%, 96.65% and 94.84% correspondingly.

Table 1: Gait classifier outcome of EAAA-HDLGR approach with other ML techniques on TD features

| **Accuracy (%) - TD Features** | | | | | | |
|---|---|---|---|---|---|---|
| **Class** | **SVM** | **ELM** | **LSTM** | **DBN** | **SSA-DBN** | **EAAA-HDLGR** |
| Pre-Stance | 95.00 | 96.07 | 93.49 | 93.63 | 96.99 | 96.07 |
| Mid-Stance | 95.15 | 94.45 | 95.89 | 92.36 | 96.49 | 95.56 |
| Terminal-Stance | 92.01 | 95.50 | 97.33 | 94.73 | 95.78 | 98.95 |
| Pre-Swing | 92.17 | 92.77 | 93.89 | 92.65 | 92.03 | 95.78 |
| Terminal-Swing | 96.92 | 94.79 | 92.76 | 96.65 | 94.84 | 98.03 |
| **Average** | **94.25** | **94.72** | **94.67** | **94.00** | **95.23** | **96.88** |

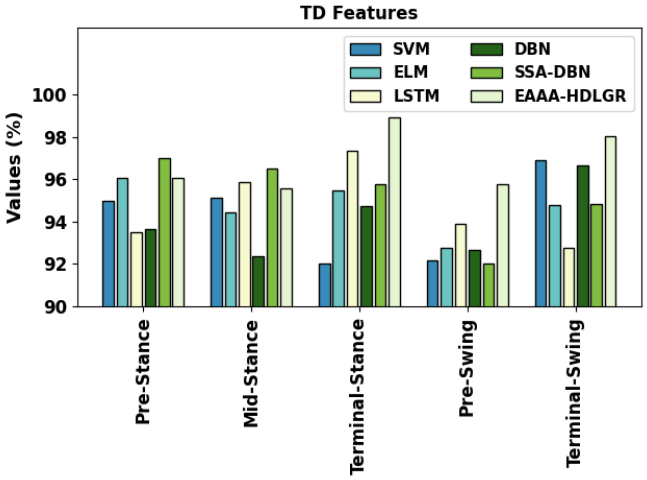



Fig. 3 Gait classifier outcome of EAAA-HDLGR approach on TD features

An average $accu_y$ assessment of the EAAA-HDLGR model with other models on TD features is given in Fig. 4. The outcomes stated that the EAAA-HDLGR method had reached improving $accu_y$ of 96.88%. Contrastingly, the SVM, ELM, LSTM, DBN, and SSA-DBN models have accomplished degrading $accu_y$ of 94.25%, 94.72%, 94.67%, 94%, and 95.23%, respectively.

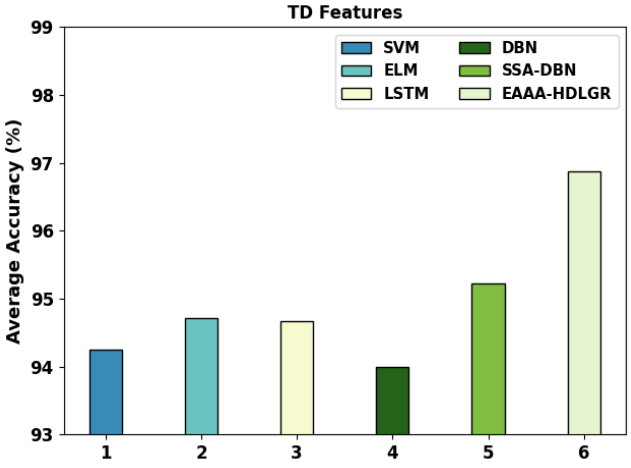



Fig. 4 Average outcome of EAAA-HDLGR approach on TD features

Tab. 2 and Fig. 5 illustrate the overall gait classification outcomes of the EAAA-HDLGR approach with other ML models on FD features. The experimental outcomes inferred that the EAAA-HDLGR system exhibits effectual outcomes. For sample, with pre-stance stage, the EAAA-HDLGR method has exhibited maximal $accu_y$ of 96.06% while the SVM, ELM, LSTM, DBN, and SSA-DBN models have attained lesser $accu_y$ of 97.02%, 92.59%, 96.62%, 97.45% and 94.59% correspondingly. In the meantime, with Terminal-Stance stage, the EAAA-HDLGR system has highlighted increasing $accu_y$ of 97.02% while the SVM, ELM, LSTM, DBN, and SSA-DBN methodologies have attained reducing $accu_y$ of 92.02%, 94.43%, 93.70%, 94.84% and 96.28% correspondingly. Lastly, with Terminal-Swing stage, the EAAA-HDLGR model has demonstrated increasing $accu_y$ of 97.44% while the SVM, ELM, LSTM, DBN, and SSA-DBN approaches have attained reducing $accu_y$ 97.45%, 95.47%, 94.38%, 92.54% and 95.29% correspondingly.

Table 2: Gait classifier outcome of EAAA-HDLGR approach with other ML techniques on FD features

| **Accuracy (%) - FD Features** | | | | | | |
|---|---|---|---|---|---|---|
| **Class** | **SVM** | **ELM** | **LSTM** | **DBN** | **SSA-DBN** | **EAAA-HDLGR** |
| Pre-Stance | 97.02 | 92.59 | 96.62 | 97.45 | 94.59 | 96.06 |
| Mid-Stance | 94.80 | 96.64 | 93.19 | 92.00 | 94.86 | 96.79 |
| Terminal-Stance | 92.02 | 94.43 | 93.70 | 94.84 | 96.28 | 97.02 |
| Pre-Swing | 96.58 | 96.98 | 95.77 | 92.96 | 97.62 | 96.55 |
| Terminal-Swing | 97.45 | 95.47 | 94.38 | 92.54 | 95.29 | 97.44 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Average** | **95.57** | **95.22** | **94.73** | **93.96** | **95.73** | **96.77** |

An average $accu_y$ analysis of the EAAA-HDLGR approach with other models on FD features are given in Fig. 6. The outcomes pointed that the EAAA-HDLGR algorithm has reached improving $accu_y$ of 96.77%. Contrastingly, the SVM, ELM, LSTM, DBN, and SSA-DBN models have accomplished degrading $accu_y$ of 95.57%, 95.22%, 94.73%, 93.96% and 95.73% correspondingly.

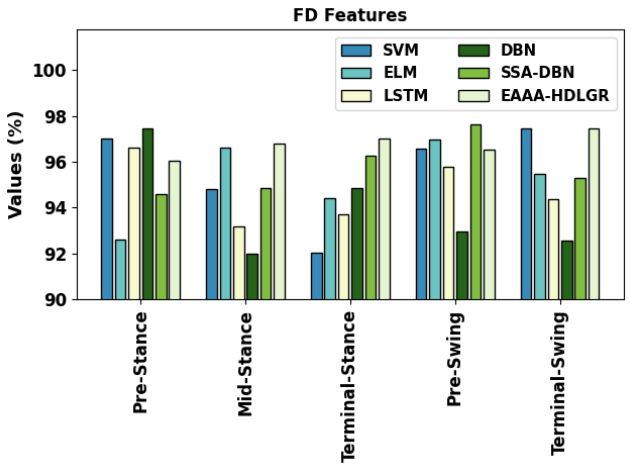



Fig. 5 Gait classifier outcome of EAAA-HDLGR approach on FD features

| **Accuracy (%) - Fusion Features** | | | | | | |
|---|---|---|---|---|---|---|
| **Class** | **SVM** | **ELM** | **LSTM** | **DBN** | **SSA-DBN** | **EAAA-HDLGR** |
| Pre-Stance | 94.53 | 94.65 | 93.41 | 97.54 | 95.24 | 97.66 |
| Mid-Stance | 92.36 | 95.09 | 93.88 | 97.12 | 96.30 | 98.23 |
| Terminal-Stance | 93.59 | 95.77 | 94.02 | 92.36 | 95.50 | 98.62 |
| Pre-Swing | 96.62 | 92.61 | 95.00 | 96.28 | 93.23 | 97.86 |
| Terminal-Swing | 92.77 | 97.00 | 94.89 | 93.75 | 97.81 | 97.55 |
| **Average** | **93.97** | **95.02** | **94.24** | **95.41** | **95.62** | **97.98** |

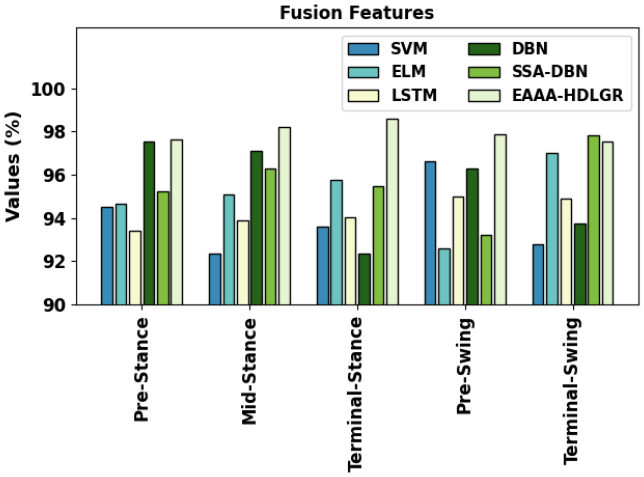



Fig. 7 Gait classifier outcome of EAAA-HDLGR approach on fusion features

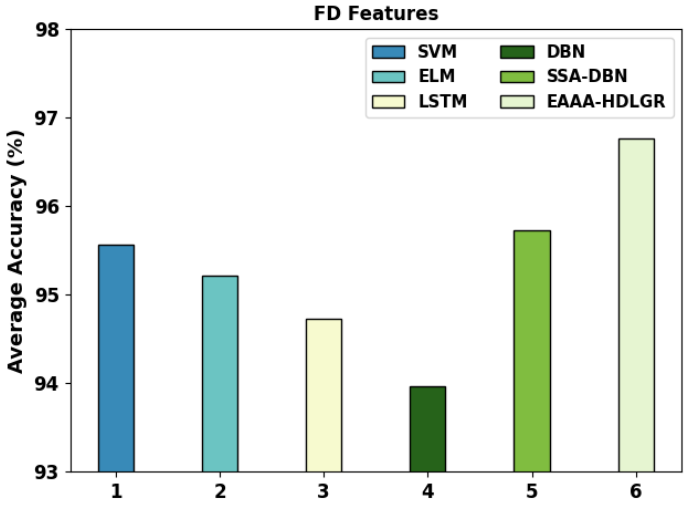



Fig. 6 Average outcome of EAAA-HDLGR approach on FD features

Tab. 3 and Fig. 7 showcase the overall gait classification outcomes of the EAAA-HDLGR algorithm with other ML techniques on Fusion features. The experimental results inferred that the EAAA-HDLGR system demonstrates effectual outcomes. For instance, with pre-stance stage, the EAAA-HDLGR approach has highlighted higher $accu_y$ of 97.66% while the SVM, ELM, LSTM, DBN, and SSA-DBN methods have gained minimal $accu_y$ of 94.53%, 94.65%, 93.41%, 97.54% and 95.24% correspondingly. Followed by, with Terminal-Stance stage, the EAAA-HDLGR algorithm has exhibited increasing $accu_y$ of 98.62% while the SVM, ELM, LSTM, DBN, and SSA-DBN models have attained reducing $accu_y$ of 93.59%, 95.77%, 94.02%, 92.36% and 95.50% respectively. Finally, with Terminal-Swing stage, the EAAA-HDLGR approach has depicted superior $accu_y$ of 97.55% while the SVM, ELM, LSTM, DBN, and SSA-DBN systems have obtained reducing $accu_y$ 92.77%, 97.00%, 94.89%, 93.75% and 97.81% correspondingly.

Table 3: Gait classifier outcome of EAAA-HDLGR approach with other ML techniques on fusion features

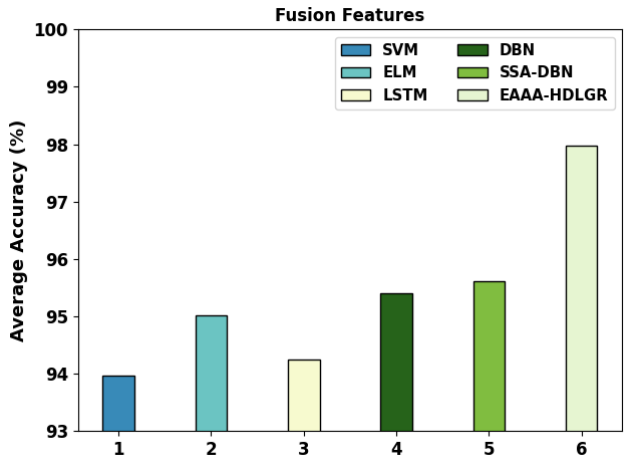



Fig. 8 Average outcome of EAAA-HDLGR approach on fusion features

An average $accu_y$ investigation of the EAAA-HDLGR approach with other methodologies on Fusion features are given in Fig. 8. The outcomes stated that the EAAA-HDLGR system has reached improving $accu_y$ of 97.98%. Also, the SVM, ELM, LSTM, DBN, and SSA-DBN systems have

accomplished degrading $accu_y$ of 93.97%, 95.02%, 94.24%, 95.41% and 95.62% correspondingly.

The TACC and VACC of the EAAA-HDLGR approach are investigated on gait classifier performance in Fig. 9. The figure stated that the EAAA-HDLGR methodology has shown higher performance with improved values of TACC and VACC. It is observable that the EAAA-HDLGR methodology has reached maximal TACC outcomes.

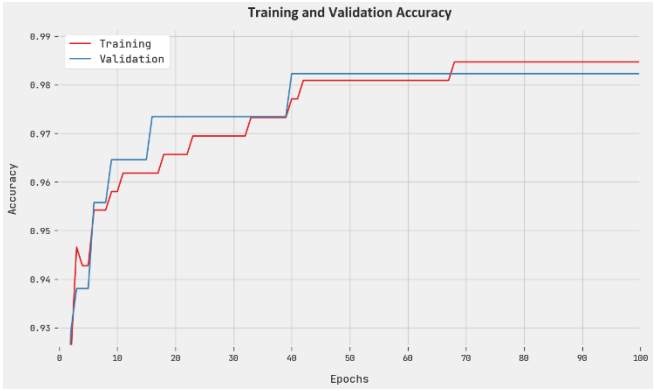



Fig. 9 TACC and VACC outcome of EAAA-HDLGR approach

The TLS and VLS of the EAAA-HDLGR system are tested on gait classifier performance in Fig. 10. The figure pointed out that the EAAA-HDLGR algorithm has better performance with minimal values of TLS and VLS. It is noticeable that the EAAA-HDLGR model has resulted to lesser VLS outcomes.

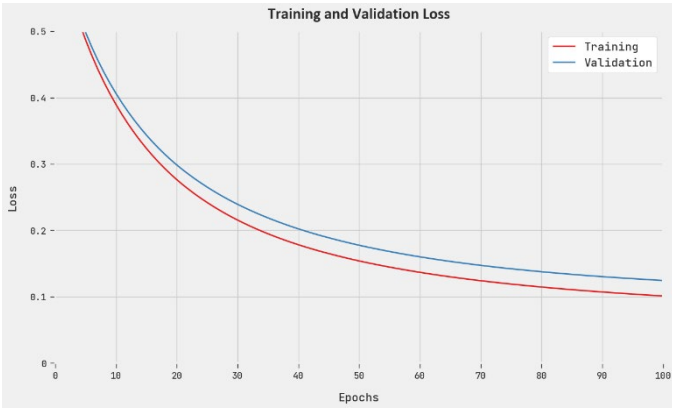



Fig. 10 TLS and VLS outcome of EAAA-HDLGR approach

## IV. Conclusion

In this study, we have derived a new EAAA-HDLGR technique for gait recognition using sEMG signals. Primarily, the EAAA-HDLGR technique derived the TD as well as FD features from the sEMG signals which are then fused together. In addition, the EAAA-HDLGR technique exploited the HDL model for gait recognition. At last, EAAA based hyperparameter optimizer is applied for the HDL model, which is mainly derived by the use of the QOBL concept. A brief classifier outcome of the EAAA-HDLGR technique is examined under diverse aspects and the results indicated the betterment of the EAAA-HDLGR technique. The results imply that the EAAA-HDLGR technique accomplishes improved results with the inclusion of EAAA on gait recognition. In future, feature reduction and feature selection processes can be combined to boost the recognition rate of the EAAA-HDLGR technique.

## Acknowlegement

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI21C1831) and the Soonchunhyang University Research Fund.

## References

[1] Z. Ni and B. Huang, "Human identification based on natural gait micro-Doppler signatures using deep transfer learning," *IET Radar, Sonar & Navigation*, vol. 14, no. 10, pp. 1640-1646, 2020.

[2] X. Bai, Y. Hui, L. Wang and F. Zhou, "Radar-based human gait recognition using dual-channel deep convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 9767-9778, 2019.

[3] M. Sharif, M. Attique, M.Z. Tahir, M. Yasmim, T. Saba *et al.* "A machine learning method with threshold based parallel feature fusion and feature selection for automated gait recognition," *Journal of Organizational and End User Computing*, vol. 32, no. 2, pp. 67-92, 2020.

[4] H. Wu, X. Zhang, J. Wu and L. Yan, "Classification algorithm for human walking gait based on multi-sensor feature fusion," in *International Conference on Mechatronics and Intelligent Robotics*, Advances in Intelligent Systems and Computing book series, Springer, Cham, vol. 856, pp. 718–725, 2019.

[5] H. Arshad, M.A. Khan, M. Sharif, M. Yasmin and M.Y. Javed, "Multi-level features fusion and selection for human gait recognition: an optimized framework of Bayesian model and binomial distribution," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 12, pp. 3601-3618, 2019.

[6] N. Mansouri, M. A. Issa and Y. B. Jemaa, "Gait features fusion for efficient automatic age classification," *IET Computer Vision*, vol. 12, no. 1, pp. 69-75, 2018.

[7] D. Thakur and S. Biswas, "Feature fusion using deep learning for smartphone based human activity recognition," *International Journal of Information Technology*, vol. 13, no. 4, pp. 1615-1624, 2021.

[8] F.M. Castro, M.J. Marín-Jiménez, N. Guil and N. Pérez de la Blanca, "Multimodal feature fusion for CNN-based gait recognition: an empirical comparison," *Neural Computing and Applications*, vol. 32, no. 17, pp. 14173-14193, 2020.

[9] M.M. Hasan and H.A. Mustafa, "Multi-level feature fusion for robust pose-based gait recognition using RNN," *International Journal of Computer Science and Information Security*, vol. 18, no. 2, pp. 20-31, 2021.

[10] Y. Lang, Q. Wang, Y. Yang, C. Hou, D. Huang *et al.* "Unsupervised domain adaptation for micro-Doppler human motion classification via feature fusion," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 3, pp. 392-396, 2018.

[11] M. Li, S. Tian, L. Sun and X. Chen, "Gait analysis for post-stroke hemiparetic patient by multi-features fusion method," *Sensors*, vol. 19, no. 7, pp. 1737, 2019.

[12] A. Abdelbaky and S. Aly, "Two-stream spatiotemporal feature fusion for human action recognition," *The Visual Computer*, vol. 37, no. 7, pp. 1821-1835, 2021.

[13] F.F. Wahid, "Statistical features from frame aggregation and differences for human gait recognition," *Multimedia Tools and Applications*, vol. 80, no. 12, pp. 18345-18364, 2021.

[14] Q. Hong, Z. Wang, J. Chen and B. Huang, "Cross-view gait recognition based on feature fusion," in *IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)*, Washington, DC, USA, pp. 640-646, 2021

[15] K. Sugandhi, F.F. Wahid and G. Raju, "Inter frame statistical feature fusion for human gait recognition," in *International Conference on Data Science and Communication (IconDSC)*, Bangalore, India, pp. 1-5, 2019

[16] M.A. Khan, H. Arshad, R. Damaševičius, A. Alqahtani, S. Alsubai *et al.* "Human gait analysis: A sequential framework of lightweight deep learning and improved moth-flame optimization algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–13, 2022.

[17] Y. Liang, E.H.K. Yeung and Y. Hu, "Parallel CNN classification for human gait identification with optimal cross data-set transfer learning," in *IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, Hong Kong, China, pp. 1-6, 2021

[18] F. Saleem, M.A. Khan, M. Alhaisoni, U. Tariq, A. Armghan *et al.* "Human gait recognition: A single stream optimal deep learning features fusion," *Sensors*, vol. 21, no. 22, pp. 7584, 2021.

[19] L. Hashem, R. Al-Harakeh and A. Cherry, "Human gait identification system based on transfer learning," in *21st International Arab*

*Conference on Information Technology (ACIT)*, Giza, Egypt, pp. 1-6, 2020.

[20] A. Mehmood, M.A. Khan, M. Sharif, S.A. Khan, M. Shaheen *et al.* "Prosperous human gait recognition: an end-to-end system based on pre-trained CNN features selection," *Multimedia Tools and Applications*, pp. 1-21, 2020, doi: 10.1007/s11042-020-08928-0

[21] M.I. Sharif, M.A. Khan, A. Alqahtani, M. Nazir, S. Alsubai *et al.* "Deep learning and kurtosis-controlled, entropy-based framework for human gait recognition using video sequences," *Electronics*, vol. 11, no. 3, pp. 334, 2022.

[22] J. He, F. Gao, J. Wang, Q. Wu, Q. Zhang *et al.* "A method combining multi-feature fusion and optimized deep belief network for emg-based human gait classification," *Mathematics*, vol. 10, no. 22, pp. 4387, 2022.

[23] R.J. Kavitha, C. Thiagarajan, P.I. Priya, A.V. Anand, E.A. Al-Ammar *et al.* "Improved harris hawks optimization with hybrid deep learning based heating and cooling load prediction on residential buildings," *Chemosphere*, vol. 309, pp. 136525, 2022.

[24] K.I. Anwer and S. Servi, "Clustering method based on artificial algae algorithm," in *International Journal of Intelligent Systems and Applications in Engineering*, vol. 9, no. 4, pp. 136-151, 2021.

[25] J. Xia, H. Zhang, R. Li, Z. Wang, Z. Cai *et al.* "Adaptive barebones salp swarm algorithm with quasi-oppositional learning for medical diagnosis systems: a comprehensive analysis," *Journal of Bionic Engineering*, vol. 19, no. 1, pp. 240-256, 2022.

# Improved Hybrid Swarm Intelligence for Optimizing the Energy in WSN

P.Kavitha Rani[1], JinHyung Kim[2], Yunyoung Nam[3,*] and Mohamed Abouhawwash[4,5]

[1]*Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore,641008, India.*
[2]*B LIFE Inc., Republic of Korea.*
[3]*Department of ICT Convergence, Soonchunhyang University, South Korea.*
[4]*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.*
[5]*Department of Computational Mathematics, Science, and Engineering (CMSE), Michigan State University, East Lansing, MI, 48824 USA.*
*Contact: ynam@sch.ac.kr

*Abstract*— **In this current century, most industries are moving towards automation where human intervention is dramatically reduced. This revolution leads to industrial revolution 4.0, which uses the Internet of Things (IoT) and wireless sensor networks (WSN). With its associated applications, this IoT device is used to compute the received WSN data from devices and transfer it to remote locations for assistance. In general, WSNs, the gateways are a long distance from the base station (BS) and are communicated through the gateways nearer to the BS. At the gateway, which is closer to the BS, energy drains faster because of the heavy load, which leads to energy issues around the BS. Since the sensors are battery-operated, either replacement or recharging of those sensor node batteries is not possible after it is deployed to their corresponding areas. In that situation, energy plays a vital role in sensor survival. Concerning reducing the network energy consumption and increasing the network lifetime, this paper proposed an efficient cluster head selection using Improved Social spider Optimization with a Rough Set (ISSRS) and routing path selection to reduce the network load using Improved Grey wolf optimization (IGWO) approach. (i) Using ISSRS, the initial clusters are formed with the local nodes, and the cluster head is chosen. (ii) Load balancing through routing path selection using IGWO. The simulation results prove that the proposed optimization-based approaches efficiently reduce the energy through load balancing compared to existing systems in terms of energy efficiency, packet delivery ratio, network throughput, and packet loss percentage.**

## I. INTRODUCTION

Wireless Sensors Network (WSN) is a collection of network devices installed in the sensing environment [1]. WSN was widely used in monitoring environments, health care, surveillance, disaster monitoring, and so on [3]. In the network, the individual devices are termed sensor nodes (SN) [2]. These sensor nodes can collect the information from the environment and transfer it into the base station (BS) directly or through intermediate gateways via wireless transmission. The transmission from the SN directly to BS causes an overhead of communication. The sensor nodes are combined as various cluster groups. One among the cluster nodes acts as a head and represents the terminal between SN and BS. This superior node is termed a cluster head (CH). The sensed data is collected from SN within its cluster and transmitted into BS. This kind of WSN is termed a cluster-based WSN.

To improve network performances, WSN has been widely used in various domains [4, 5]. The primary reasons behind using sensors in the environment are their easy configuration and management setup [6]. These sensor nodes are autonomously operated and compute their network infrastructure ad hoc. In this kind of scenario, the nodes are unstable, and they can connect to their Neighbour based on certain factors for data transmission. The CH combines the sensed data and relays it to the BS. CH can efficiently construct its single-hop and multi-hop path toward BS. The end user can use the centralized base station through the internet to retrieve the required data. During data transmission, the deployed SN can be mobile or static. The static SN is referred to as non-adaptive, and its routing mechanism is unchanged. Various mobile SN are dynamic, and their routing tables are updated frequently when there is a change in network topology. While compared to dynamic routing, the static mechanism is secure; moreover, the stationary solutions are not appropriate for the

scalable network in more significant regions [7]. The communication efficiency is increased through load balance, resource utilization, and throughput by integrating the Internet of Things (IoT) in all fields [8-10]. More physical devices are involved in IoT to transmit the data using the internet. The WSN helps IoT and supports observing and forwarding the data in the physical networking environment [11]. This research article proposed an effective cluster head selection and routing approach with load balancing to reduce network energy and computation. The significant contribution of this research is as follows:

For Each cluster, the cluster head is selected using Improved Social Spider Optimization with a Roughset model (ISSRS)

- Once the cluster head for each cluster is selected, the routing from source to destination is selected using IGWO which is used as a load balancing approach to reduce the load of base station and gateways.
- The proposed cluster head selection and load balancing are evaluated using simulation by calculating the total number of ko nodes that are dead and alive. Further energy efficiency, throughput, and execution cost are calculated.

## II. RELATED WORK

The section discusses the existing literature on loading balancing, routing, and clustering in WSN-based IoT systems. Majid et al., [12] did a systematic review of WSN and IoT-based applications. The detailed discussion using WSN in industry 4.0 with its IoT-related applications, research gaps, and future research directions are discussed. Lipare et al., [13] developed GWO with a novel fitness function-based load-balancing approach to overcome the issues of energy holes in WSN. In this work, authors applied GWO for routing and cluster with enhanced fitness functions. Kumar et al. [14] developed a clustering with a Centroid-oriented routing protocol for WSN-based IoT. This routing approach has three phases: network initiation to create the zone among the nodes, selection of zone coordinator, and zone head selection. The path from the zone head to the base station is optimized and distance centric with one or dual-hop level transmission to reduce packet loss.

Kashif et al. [15] used a gateway node to reduce the load of the cluster head and transmit the data to the base station. Their assumption that the gateway node should be in the Neighbour gateway range leads to high complexity and also limits the sensor node deployment randomly. Arroyo et al., [16] developed a cloud computing and WSN-based air quality monitoring system. The distributed sensors are connected to the cloud for storage, monitoring, and processing from WSN. The data processing is performed using artificial intelligence approaches. Mellado et al., [17] predict the shadow of satellites using WSN and IoT for smart cities. The authors developed a graphical method for predicting satellite coverage in urban areas. Their captured shadow images are similar to Eutelsat satellite measurements.

Haseeb et al. [18] proposed a secure IoT with a WSN-based innovative agriculture application. Initially, the agricultural data are sensed using the sensors, and cluster heads are formed using the multi-criteria decision method. The signal strength is measured using the Signal Noise Ratio (SNR). The simulated results prove that the proposed model enhances the network performance with increased throughput and reduced drop ratio.

Butun et al., [19] reviewed the challenges and vulnerabilities of IoT and WSN-based applications. Further authors also discussed attacks related to IoT and WSN with its solutions and advantages with limitations.

Pamarthi et al., [20] did a systematic review of the WSN security issues for IoT applications. Pamarthi et al. discussed routing attacks with their security measures for ad hoc networks. Parvathy [21] discussed the wormhole attacks related to WSN and IoT. The authors examined the interconnection between IoT and WSN through the internet and it is guarded against external aggression. Raza et al., [22] reviewed the applications and security issues for sensor networks using the MAC protocol. The security challenges of various layers with their respective solutions were presented.

The technological security issues, with their advantages and challenges, are profoundly concerning. Kumar et al. [23] discussed IoT applications with evolutionary approaches. Sharma et al., [24] designed a relay node structure or WSN and load balancing. Here WSN created an IoT with a WSN-based application to monitor the natural disaster in the coal mine. WSN uses different parameters for network analysis, including the network's lifetime, energy consumption, throughput, and area coverage.

Faheem et al. [25] developed an approach using bio-inspired for WSN-assisted innovative application of grid. This application developed an IoT, and WSN-assisted farming robot using a computer vision-based approach is introduced. It has been used to classify weed and non-image. The weed image data are detected through WSN, and mutual authentication protocols for WSN-based IoT scenarios are also done. This scheme has improved performance, low communication, and execution costs. The main issue is the security concern for heterogeneous IoT environments needs to be considered.

## III. METHODOLOGY

This section proposed efficient cluster head selection and routing path selection for load balancing in WSN assisted IoT environment.

### *A. System Model*

The proposed system overview is shown in Fig. 1. The sensor node (SN) collects the data from the IoT sensors, such as temperature, pressure, and so on, concerning the applications. The gathered data are converted into digital form, which is transferred into the base station (BS) through cluster heads in the wireless networks. The BS consists of more power and memory. Further, it is connected to the best resources to save energy. In IoT-based applications, BS can store the data, analyze it and visualize data collected from the cluster head. BS can also provide a graphical user interface for the user to direct interaction or forward the collected data to the remotely managed server through the internet. The remoter servers are responsible for sending the sensed input data to the authorized user. This data is also saved on web pages to access worldwide through the internet. In this work, the clusters are formed using Euclidean distance, and the cluster head for each group is selected using the ISSRS approach. In the host computer, the

load is balanced using IGWO to reduce system energy. The end users can access the centralized base station through the internet to retrieve the required data.

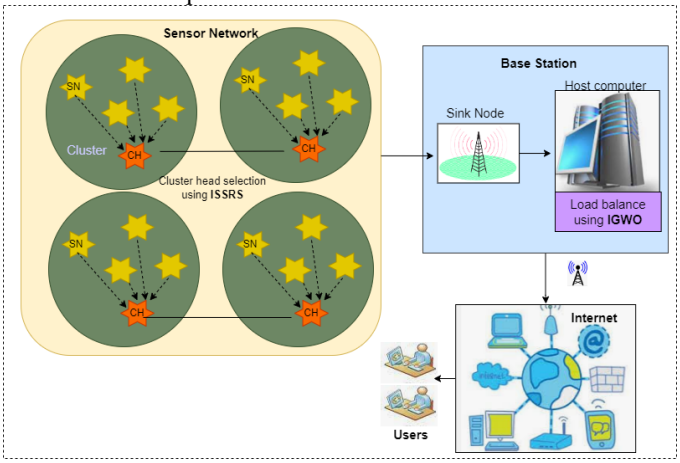



Fig. 1 Architecture of proposed load balancing in WSN-assisted IoT model

## B. Cluster Formation

Each SN is allotted its identification number. The links are formed from the source SN to the destination using Euclidean distance. It measures the distance between two adjacent SN located within the communication range CR. The sensing range SNR has been denoted as the range between the SN that can sense the signal. N is the number of nodes in the network. The set of possible communication is defined in Eq. (1)

$$C = \{(SN_i, SN_j) \in N | SN_j \; recieve \; message \; from \; SN_i\} \quad (1)$$

The Neighbouring of $SN_i$ is declared as in Eq. (2)

$$Neighbor(SN_i) = \{SN_j \in N | (SN_i, SN_j) \in l\} \quad (2)$$

The link between the SN is limited to its capacity called $\mu_{ij}$

$$P_{ij} = \sum_C X_{ij}^C \leq \mu_{ij} \quad (3)$$

Where between$X_{ij}^C$ - set of communication link betwen$(i,j) \in L$.

## C. Cluster Head Selection Using ISSRS

In Various researchers have proposed cluster-based concepts to simplify network management. The cluster head node can able to manage the sensor node. It is responsible for arranging the SN within the cluster from the routing table. Also, the cluster head gathers the data and aggregates and retransfers the collected data to the collections. These cluster heads lack energy due to the unexpected load assigned to them. To minimize the energy in CH, the distance between CH and the base station should be minimum, or else there is a need for multihop transfer from CH to the base station through intermediate cluster heads. The previous cluster method selects the cluster head node randomly without considering the parameters related to the centralized approach. This will affect the network's scalability. The proposed meta-heuristic-based load balancing and cluster head selection approach efficiently chose the CH to transmit the data to neighbor CH based on the parameters such as Initial energy, Residual energy, Energy consumption rate, and Average network energy. The significant parameter in choosing the CH is the initial energy. Residual energy means after the series of execution, and the CH considers the remaining SN energy. The energy consumption rate is the primary factor for each cluster. The average energy of the network is the reference energy, and it is the ideal energy to make the network alive. This starts with the random initialization of the spider population. Each individual is converted into the binary vector of length N using Eq. (4)

$$V = \frac{1}{1+e^{-X_i^j(t)}} \quad (4)$$

$$X_i^j(t+1) = \begin{cases} 0 & if \; V\left(X_i^j(t)\right) > \varepsilon \\ 1 & otherwise \end{cases} \quad (5)$$

Where, $X_i^j(t)$- Spider (SN) value at iteration t and $\varepsilon \in [0,1]$. It uses the dependency degree as in Eq. (6)

$$\gamma_E(D) = \frac{P_E(D)}{|N|} \quad (6)$$

Where, E is the set of conditions and D is the selected nodes, P (D) is the positive region which consists of all the nodes N that is classified based on the information C. The fitness function is declared as in Eq. (7)

$$F(S) = \tau\gamma_S(D) + (1-\tau)(1 - \frac{|S|}{|C|}) \quad (7)$$

Where, $\tau \in [0,1]$ provides the distance between number of selected nodes and classifying results. |.| represents the number of selected nodes. Each spider fitness is compared with global best ($F_{gbest}$). If it has highest fitness value then $F_{gbest}$ is replaced with the recent spider and its position which will added to the selected set S. the weight of each spider is updated using the Eq. (8)

$$w_{ij} = \frac{F_i - worst}{best - worst} \quad (8)$$

Where, F – fitness of ith spider, worst –worst fitness value of population N and best is the best fitness value of population N. In social spider the information exchange between the spiders is represented by vibration of ith spider perceived from jth spider is represented in Eq. (9)

$$vb_{i,j} = w_j e^{d_{i,j}^2} \quad (9)$$

Where, w –jth spider weight, d- distance measured between I and jth spider. There are three categories of vibration produced by the spider such as n,b, and f. where vbn is the vibration produced by the nearest spider, vbf is the vibration produced by the female spider, and vbb is the vibration produced by the male spider as the best spider. The new position of Female Spider Position (FSP) is updated using Eq. (10), which is managed by the probability factor P, and the moving factor is updated with the relation between other spider and their respective vibration in the search space.

$$FSP_i^{k+1} = \begin{cases} FSP_i^k + \alpha.vb_{i,n}.\left(s_n - FSP_i^k\right) + \beta.vb_{i,b} \\ .\left(s_b - FSP_i^k\right) + \delta.\left(r - \frac{1}{2}\right) with \; probablity \; P \\ FSP_i^k - \alpha.vb_{i,n}.\left(s_n - FSP_i^k\right) - \beta.vb_{i,b} \\ .\left(s_b - FSP_i^k\right) + \delta.\left(r - \frac{1}{2}\right) with \; probablity \; 1-P \end{cases} \quad (10)$$

Where, $\alpha, \beta, \delta \; and \; r$ - random number between [0,1], k-iteration number and $s_n$ and $s_b$ nearest spider and best spider in the web respectively. The male spider process is operated using the Eq. (11). Until the stopping condition met, this process is repeated. The final selected cluster heads are represented in S. The workflow of ISSRS is show in Fig. 2.

$$MSP_i^{k+1} = \begin{cases} MSP_i^k + \alpha . vb_{i,f} . \left(s_f - MSP_i^k\right) + \delta . \left(r - \frac{1}{2}\right) \\ \qquad if MSP_i^k \in Dominant\ group \\ SP_i^k + \alpha . \left(\frac{\sum_{h \in nondominant} MSP_h^k . w_h}{\sum_{h \in nondominant} w_h}\right) - MSP_i^k \\ \qquad if\ MSP_i^k \in Non - Dominant\ group \end{cases} \quad (11)$$

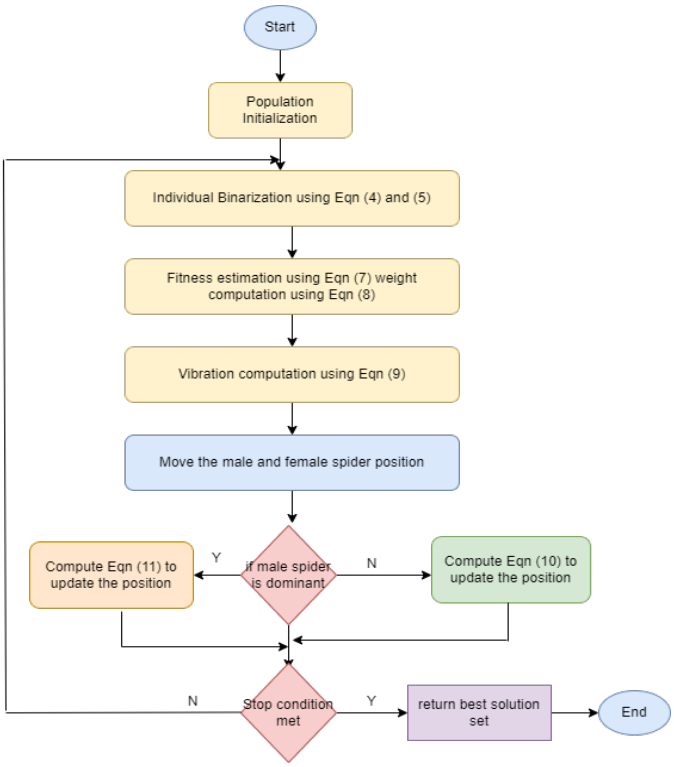



Fig. 2 Workflow of ISSRS

### D. *Routing Using IGWO for Load Balancing and Energy Saving*

Once the cluster head has been selected using ISSRS, load balancing is the next consideration through the proper routing mechanism. Since the CH is nearer the base station, it can directly transmit the sensed data to BS. In the case of CH far from the BS, it can choose multi-hop transmission using routing methods to select the efficient path to the next CH for information transfer. Assume the sensor node (SN) and the deployed gateways in the region with the area of X*X m2. Within the communication range, the SNi can connect only one Cluster head CHi. Let kij be the Boolean variable defined in Eq. (12). Each CH should secure to BS directly or through another CH via gateways. The load Li is reduced for the CH and balanced with other CH.

$$= \begin{cases} k_{ij} = 1 \; if SN_i \; connected \; to \; CH_i \\ k_{ij} = 0 \qquad\qquad otherwise \end{cases} \quad (12)$$

The Improved Grey wolf optimization (IGWO) based routing scheme consists of three phases such as (i) Wolves initialization (ii) For each wolf fitness value is computed and (iii) Velocity and Wolves position updating.

(i) **Wolves Initialization**: Each solution is stated as a map from one CH to other CH/BS. The solution size is similar to the number of CH called 'H'. The solution represents the route from a CH towards BS via next consecutive CH in network. Each CH is stated with the random number $\left(X_{i,c}\right) = Rand\ [0,1] where\ 1 \leq i \leq N_{\ is} and\ 1 \leq c \leq H$. Where $N_{\ is}$ - number of initial solutions and c – cluster head number of corresponding solution. It maps the $CH_k$ as next Neighbour CH. The routing path towards the BS from $CH_c$ which indicates that $CH_k$ sends the data to $CH_c$. This mapping is computed using the Eq. (13)

$$CH_k = Index(SetNextCH(CH_c), n) \quad (13)$$

Where, this indexing function indicates the index of nth CH and $n = Ceil((X_{i.c}) \times |SetNextCH(CH_c)|)$.

(ii) **Fitness Value Computation:** The fitness function is the important factor that measure the quality of the result based on the parameters. This phase helps to update the parameters such as $\alpha, \beta \; and \; \delta$ values at each iteration. The novelty of this paper is updated here to set the fitness value to generate the routing path from each CH to BS. The entire distance (ED) travelled from CH to BS is represented in Eq. (14)

$$ED = \sum_{i=1}^{h} dist(CH_i, NextCH(CH_i)) \quad (14)$$

Where, h- number of CH. The total hops from CH in the network is defined in Eq. (15)

$$HP = \sum_{i=1}^{h} NextCount(CH_i) \quad (15)$$

Routing is performed with the consideration of minimum distance traveled with less number of hops. Hence, the smaller the space, the Hops traversed leads high fitness value for obtaining the best solution. Therefore, the fitness value is inversely proportional to the number of hops and distance. The highest fitness values lead to the best solution. The Routing fitness (RF) function is represented in Eq. (16)

$$RF = \frac{K_1}{(w_1 \times ED + w_2 \times HP)} \quad (16)$$

Where, $K_1$ – proportionality constant, $(w_1, w_2) \in [0,1]$ and $w_1 + w_2 = 1$ [13]. This RF can balance the number of hops and distance from one CH to other towards BS.

(iii) **Velocity and Position Updation**: Each wolf of $\alpha, \beta \; and \; \delta$ (solutions) needs to update the location depending on the positions using the Eq. (17) in order to reach the prey (best feasible solution) [30].

$$D_\alpha = |C_1 . X_\alpha - X| \quad (17)$$

$$D_\beta = |C_2.X_\beta - X| \quad (18)$$

$$D_\delta = |C_3.X_\delta - X| \quad (19)$$

Where, X – wolf current position, $X_\alpha$- current position of $\alpha$ wolf, $D_\alpha$- updated position of $\alpha$ wolf, $X_\beta$ and $X_\delta$ are the current position of $\beta$ and $\delta$ wolfs respectively. And $D_\beta$ and $D_\delta$ are the updated position of $\beta$ and $\delta$ wolfs respectively. In GWO [4], the alpha wolf is the global solution, the best solution from the most recent iteration is represented by delta wolf, while the best solution from the previous iteration is represented by beta wolf. The position of each wolf is updated using the Eq. (20)

$$X_{t+1} = \frac{X_1+X_2+X_3}{3} \quad (20)$$

$$X_1 = |X_\alpha - R_1.D_\alpha| \quad (21$$

$$X_2 = |X_\beta - R_2.D_\beta| \quad (22)$$

$$X_3 = |X_\delta - R_3.D_\delta| \quad (23)$$

Where t – current iteration and R- random vectors. However, $(X_{i,c})$ is in the range [0,1]. Once the new positions are assigned, the solutions are re-evaluated. Fig. 3 represents the routing path from CH to another CH towards BS. CH2 and CH4 are nearer to BS and they can send the sensed data directly to BS. Rather, CH1 and CH 3 are far away from the BS and it is communicated to BS through multi hop transmission where, CH1 send data through CH2 (nearer to CH1 than CH4) towards BS and CH3 send data through CH4 (nearer to CH3 than CH2) towards BS. This efficient routing approach reduces the energy of the system.

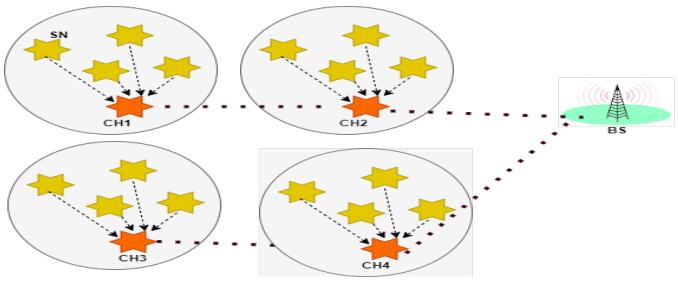



Fig. 3 Proposed IGWO based routing

Fig.4. represents the Cluster head-based load balancing approach using proposed ISSRS-IGWO mechanism. Initially, CH1 senses the data from five SN, rather CH3 senses data from three SN. CH1 have over workload and it is balanced with CH3. As well as, CH2 consists of four SN and CH4 consists of two SN. The load of CH2 is balanced with CH4. This efficient load balancing reduces the computational overhead of the network.

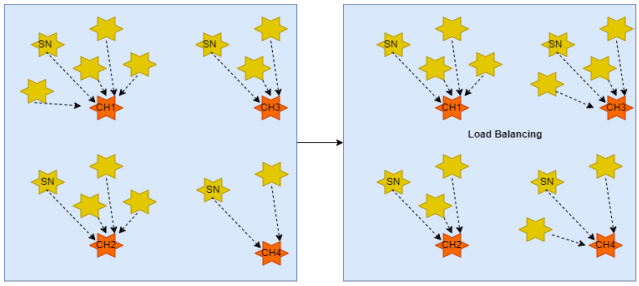



Fig. 4 Load balancing of CH using Proposed ISSRS-IGWO

### *E. Energy Model*

In order to transmit one bit of data over the distance d, the energy is computed as in Eq. (24)

$$E_{CX}(k,d) = \begin{cases} k \times E_{cr} + k \times E_{cf} \times d^2, & if\ d < \sqrt{\frac{E_{cf}}{E_p}} \\ k \times E_{cr} + k \times E_p \times d^4, & if\ d \geq \sqrt{\frac{E_{cf}}{E_p}} \end{cases} \quad (24)$$

Where, $E_{cr}$- Energy to transmit or receive a bit for k nodes, $E_{cf}$- free space amplifier energy, $E_p$- multi path amplifier energy. The energy for receiving bit is computed as in Eq. (25)

$$E_R(K) = k \times E_{cr} \quad (25)$$

## IV. Simulation Results and Discussions

The proposed system model performance is compared with standard Grey wolf optimization, Multi-objective fuzzy clustering algorithm (MOFCA), and Particle Swarm Optimization (PSO) - ECHS [26]. The evaluations regarding the number of dead and alive nodes, load balancing, energy consumption, and packet loss are implemented using Python Scikit learn [27-34].

MOFCA approach is made by using both clustering and routing for multi-hop transmission. The proposed approach uses two methods for its efficient routing mechanism. The cluster head selection is based on ISSRS, and the routing is based on IGWO. PSO used only one best possible solution. At the same time, the proposed model used three possible explanations. The routing fitness value comparison of IGWO and other approaches is shown in Tab. 2 for equal and unequal numbers of nodes. From Tab.1. the proposed ISSRS-IGWO secured improved routing fitness values compared to different methods for the equal and unequal load. Increasing the fitness values gives a better solution.

TABLE I. ROUTING FITNESS VALUES OF PROPOSED ISSRS-IGWO VS EXISTING MODELS

| Load of SN | Equal loads | | | Unequal loads | | |
|---|---|---|---|---|---|---|
| No. of SN | 100 | 300 | 500 | 100 | 300 | 500 |
| GWO | 1.31 | 1.46 | 1.71 | 0.81 | 0.92 | 0.11 |
| MOFCA | 1.52 | 1.68 | 1.86 | 1.14 | 1.34 | 1.64 |
| PSO-ECHS | 1.42 | 1.53 | 1.61 | 0.91 | 1.2 | 1.7 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Proposed ISSRS-IGWO | 1.72 | 1.81 | 1.98 | 1.45 | 1.52 | 1.79 |

Evaluation of the number of alive nodes during the transmission is shown in Fig. 5 for the number of rounds. From the illustration, it has been observed that the proposed model secured longer network life than other approaches. For 3000 matches, using the proposed ISSRS-IGWO model,850 nodes are alive for transmission. At the same time, in the different approaches, such as GWO, MOFCA, and PSO-ECHS, the number of active nodes is 630, 580, and 610, respectively. Hence, the proposed model increases the network lifespan while transmitting the data.

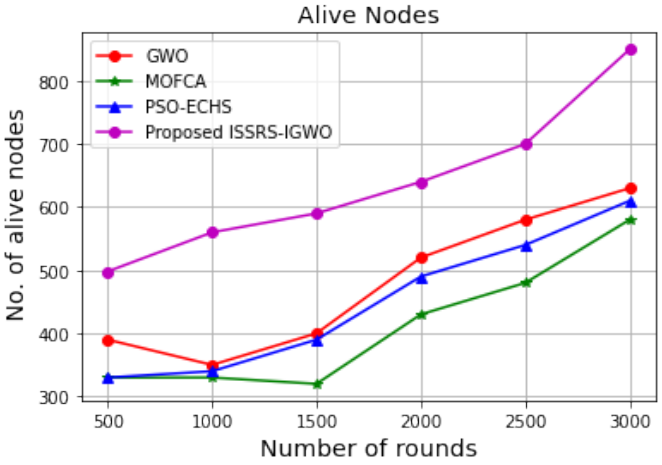



Fig. 5 Number of Alive nodes for proposed ISSRS-IGWO vs existing approaches

Fig. 6 displays the evaluation in terms of the number of dead nodes. It can be seen from the illustration that the proposed model secured fewer dead nodes than other methods. For 3000 rounds, 330 nodes died during the transmission while using the proposed ISSRS-IGWO model. Unlike the different approaches, such as GWO and MOFCA, the number of dead nodes is 410 500, and for PSO-ECHS, all 500 nodes have died in 2500 rounds. This result reveals that the proposed algorithm optimizes the energy consumption of nodes, so that cluster heads are rotating among nodes. This helps to minimize energy consumption by avoiding new nodes as cluster heads in the network. Hence, the proposed model increases the network lifespan with the reduced number of dead nodes while transmitting the data.

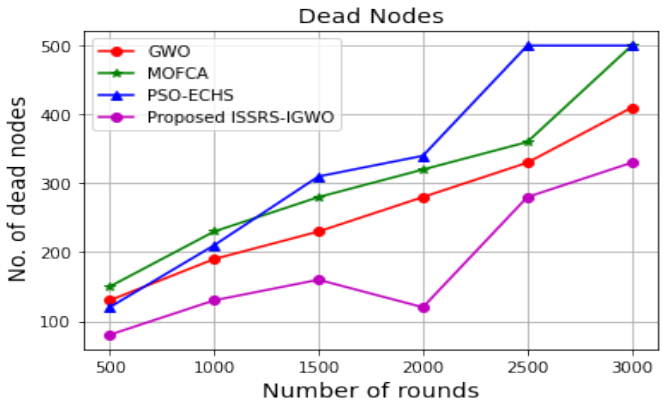



Fig. 6 Number of dead nodes for proposed ISSRS-IGWO vs existing approaches

The number of data packets transferred from CH to BS is shown in Fig.7. in terms of the number of rounds. This illustration shows the superior performance of the proposed model to send the sensed data from SN to BS through CH. Compared to other approaches, the efficient cluster head selection with the routing approach transmits the maximum amount of data packets to BS via CH.


Fig. 7 Number of packets transmission of proposed ISSRS-IGWO vs existing approaches

Fig. 8 compares existing and proposed load-balancing approaches in terms of throughput. Throughput is measured as the size of the number of packets delivered, and it is measured using Eq. (26). Compared to the traditional approaches, the proposed model secured increased throughput and returned the packages without loss. The increased throughput of the proposed model is due to the implementation of robust cluster head selection.

$$throughput\ \left(\frac{bit}{sec}\right) = \frac{successfully\ recieved\ packet\ count\ at\ BS \times size\ of\ packet}{total\ time\ to\ send\ the\ packet\ from\ SN\ to\ BS} \quad (26)$$

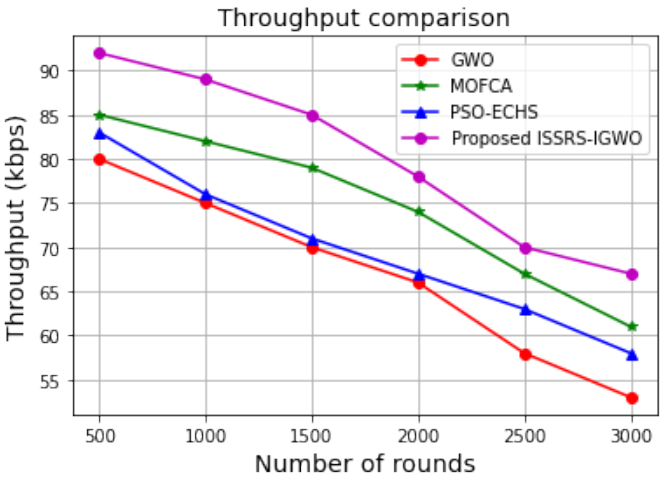



Fig. 8 Throughput comparison of existing and proposed ISSRS-IGWO models

Packet loss is computed to assess the dependability of the suggested communication line. Compared to the total number of packets detected by SN in the network, the packet loss (PL) % is similar to the not received packet by the BS. PL occurs because of the SN failure, which has the data over the sensing network area. The PL computation is shown in Eq. (28). The evaluation of the proposed model in terms of PL is demonstrated in Fig.9. Compared to the existing approaches, the proposed model secured less PL percentage at the simulation time of 200 seconds. The PL for the proposed model is 2.2%, whereas the other approaches, such as GWO, MOFCA, and PSO, 21.1%, 36%, and 40.6%, respectively.

$$PL = (Total\ packet\ sensed\ by\ SN\ at\ simulation\ time) - (total\ packet\ reicevd\ by\ BS\ at\ simulation\ time) \quad (27)$$

$$PL(\%) = \frac{PL \times 100}{packets\ sensed\ by\ SN\ at\ simulation\ time\ t} \quad (28)$$

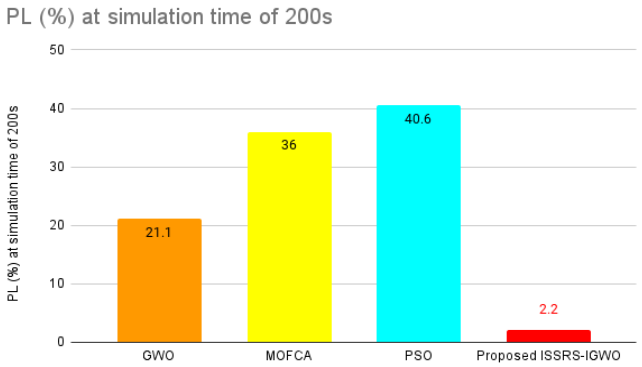



Fig. 9 PL (%) comparison of proposed ISSRS-IGWO vs existing models

The energy efficiency of the proposed model is evaluated, and it is compared with the existing approaches. The estimated results are shown in Tab. 3. The energy computation evaluation is performed on an equal and unequal number of SN for 100, 300, and 500 nodes. From Tab.1, the proposed ISSRS-IGWO consumes less energy than another meta-heuristic approaches-based load balancing. The proposed model reduces the load of SN through CH, which distributes the load to other clusters equally. The proposed routing approach chose the best path with a minimum distance from the SN to BS through CH, which will reduce the energy with the minimum distance transmission. This evaluation ensures that the proposed model consumes minimum power compared to the other approaches and increases the network lifetime.

Table II: Total energy consumption (J) of the sensor nodes using proposed ISSRS-IGWO vs existing models

| Sensor node load | Equal number of Lodes | | | Unequal number of Lodes | | |
|---|---|---|---|---|---|---|
| Number of sensor nodes | 100 | 300 | 500 | 100 | 300 | 500 |
| PSO-ECHS | 94.2 | 178.2 | 250.3 | 95.2 | 178.5 | 280.1 |
| MOFCA | 88.1 | 165.6 | 231.9 | 89.3 | 160.2 | 233.1 |
| GWO | 75.3 | 128.8 | 183.2 | 75.4 | 129.4 | 184.7 |
| Proposed ISSRS-IGWO | 52.4 | 86.6 | 112.3 | 54.2 | 87.9 | 114.3 |

Fig.10 represents the comparative analysis of proposed and existing approaches regarding execution cost. For 500 nodes, the proposed model secured 35 seconds for its execution. At the same time, the other methods, such as GWO, MOFCA, and PSO-ECHS, obtained the 80s, 95s, and 90, respectively. For this reason, the proposed ISSRS-IGWO is superior to other approaches with a minimum execution cost. Hence, the simulated evaluation and its comparative analysis show that the proposed model is efficient in reducing the energy and computation cost of the WSN.

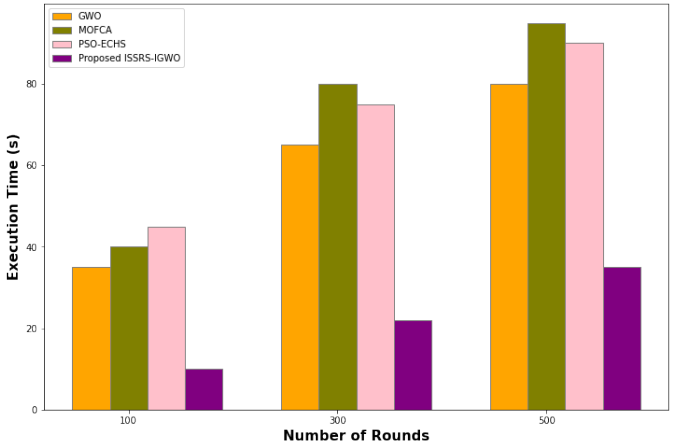



Fig. 10 Execution Time comparison

## V. Conclusions

This paper proposes an efficient cluster head selection and load-balancing approach using Swarm Intelligence for a WSN-assisted IoT environment. Initially, the cluster of sensor nodes is formed using the Euclidean distance between the Neighbour nodes in the communication range. Using the proposed Improved social spider optimization with a rough set approach, the CH is selected for each cluster, representing the manager for each group. The data from the sensor nodes are

forwarded to the base station through cluster heads. The set heads nearer to the base station are sent the sensed data directly. Instead, the clusters far from the base station drain their energy and need extra computation costs to send the data. To reduce the system energy, load, and computation cost, another swarm approach called improved grey wolf optimization had been proposed for efficient routing from SN to BS via CH. The proposed cluster head selection and load balancing were evaluated in the simulation environment regarding the number of dead and alive nodes, energy efficiency, throughput, packet loss, and execution cost. These simulation results are compared with the existing approaches. The results prove the efficiency of the proposed model with reduced energy and execution costs, which will improve the network lifetime. The sensor node mobility is not considered in this work which is regarded as the limitation. In the future, sensor nodes' mobility, cluster head, and base station can be assumed to determine the location change and provides efficient routing with reduced load and energy.

## ACKNOWLEDGMENT

This work was supported by the Collabo R&D between Industry, Academy, and Research Institute(S3250534) funded by the Ministry of SMEs and Startups (MSS, Korea), and the Soonchunhyang University Research Fund.

## REFERENCES

[1]J.Li, L.L.Andrew, C.H.Foh, M.Zukerman and H.H.Chen, "Connectivity coverage and placement in wireless sensor networks,"*Sensors*, vol.9,no.10, pp.7664-7693,2009.

[2]P.Kuila and P.K.Jana, "Improved load balanced clustering algorithm for wireless sensor networks,"in *Proc.International conference on advanced computing, networking and security,* Surathkal, India, pp. 399-404,2011.

[3]J.Tang, B.Hao and A.Sen "Relay node placement in large scale wireless sensor networks,"*Computer Communications*, vol.29,no.4, pp.490-501,2006.

[4]P.S.Mehra, M.N.Doja and B.Alam, "Fuzzy based enhanced cluster head selection (FBECS) for WSN,"*Journal of King Saud University Science*, vol.32,no.1, pp.390-401,2020.

[5]A.Tripathi, H.P.Gupta, T.Dutta, R.Mishra, K.K.Shukla *et al*.,"Coverage and connectivity in WSNs: A survey, research issues and challenges,"*IEEE Access*, vol.6, no. 22, pp.26971-26992,2018,

[6]K.A.Awan, I.U.Din, A.Almogren, M.Guizani and S.Khan, "Stab Trust A stable and centralized trustbased clustering mechanism for IoT enabled vehicular ad-hoc networks," *IEEE Access*, vol.8, no. 4, pp.21159-21177,2020.

[7]A.Abuarqoub,M.Hammoudeh, B.Adebisi, S.Jabbar, A.Bounceur *et al*., "Dynamic clustering and management of mobile wireless sensor networks,"*Computer Networks*, vol.117, no. 11, pp.62-75,2017.

[8]H.A.Khattak, Z.Ameer, U.I.Din and M.K.Khan, "Cross-layer design and optimization techniques in wireless multimedia sensor networks for smart cities,"*Computer Science and Information Systems*, vol.16,no.1, pp.1-17,2019.

[9]I.U.Din,M.Guizani,S.Hassan,B.S.Kim,M.K.Khan *et al*.,"The internet of things: A review of enabled technologies and future challenges," *IEEE Access*, vol.7, no. 4, pp.7606-7640,2018.

[10] K.A.Awan, I.U.Din, A.Almogren, M.Guizani, A.Altameem *et al*.,"Robust trust pro-privacy robust distributed trust management mechanism for internet of things,"*IEEE Access*, vol.7, no. 9, pp.62095-62106,2019.

[11] I.U.Din, M.Guizani, J.J.Rodrigues, S.Hassan and V.V.Korotaev, "Machine learning in the internet of things: Designed techniques for smart cities,"*Future Generation Computer Systems*, vol.100, no.5, pp.826-843,2019.

[12] M.Majid, S.Habib, A.R.Javed, M.Rizwan, G.Srivastava *et al*.,"Applications of wireless sensor networks and internet of things frameworks in the industry revolution 4.0: A systematic literature review,"*Sensors*, vol.22,no.6, pp.2087, 2022.

[13] A.Lipare, D.R.Edla and V.Kuppil, "Energy efficient load balancing approach for avoiding energy hole problem in WSN using grey wolf optimizer with novel fitness function,"*Applied Soft Computing*, vol.84, no. 3, pp.105706,2019.

[14] N.P.R. Kumar and J.B. Gnanadhas, "Cluster centroid based energy efficient routing protocol for WSN assisted IoT," *Advances in Science, Technology and Engineering Systems Journal*, vol.5, no.4, pp. 296-313,2020.

[15] K.N.Qureshi, M.U.Bashir, J.Lloret and A.Le, "Optimized clusterbased dynamic energy aware routing protocol for wireless sensor networks in agriculture precision,"*Journal of Sensors*, vol.2020, no. 4, pp.1-19,2020.

[16] P.Arroyo, J.L.Herrero, J.L.Suárez and J.Loza,"Wireless sensor network combined with cloud computing for air quality monitoring,"*Sensors*, vol.19,no.3, pp. 25-45, 2019.

[17] S.Hornillo,R.Martín and V.Baena, "Prediction of satellite shadowing in smart cities with application to IoT,"*Sensors*, vol.20,no.2, pp.475-487,2020.

[18] K.Haseeb, I.Ud Din, A.Almogren and N.Islam, "An energy efficient and secure IoT-based WSN framework: An application to smart agriculture,"*Sensors*, vol.20,no.7, pp.2081-2095,2020.

[19] I.Butun,P.Österberg and H.So, "Security of the Internet of Things: Vulnerabilities, attacks, and countermeasures,"*IEEE Communications Surveys & Tutorials*, vol.22,no.1, pp.616-644,2019.

[20] S.Pamarthi and R.Narmadha, "Literature review on network security in wireless mobile ad-hoc network for IoT applications: Network attacks and detection mechanisms,"*International Journal of Intelligent Unmanned Systems*, vol. 6, no. 2, pp. 1-22, 2021.

[21] M.U.Farooq,M.Waseem, S.Mazhar,A.Khairi and T.Kamal, "A review on internet of things (IoT),"*International Journal of Computer Applications*, vol.113,no.1, pp.1-7,2015.

[22] S.Raza, M.Faheem and M.Guenes, "Industrial wireless sensor and actuator networks in industry 4.0: Exploring requirements, protocols, and challenges—A MAC survey,"*International Journal of Communication Systems*, vol.32,no.15, pp.e4074,2019.

[23] S.Kumar, P.Tiwari and M.Zymbler, "Internet of Things is a revolutionary approach for future technology enhancement: a review,"*Journal of Big Data*, vol.6,no.1, pp.1-21,2019.

[24] R.Sharma and S.Prakash, "Enhancement of relay nodes communication approach in WSNIoT for underground coal mine,"*Journal of Information and Optimization Sciences*, vol.41,no.2, pp.521-531,2020.

[25] M.Faheem, R.A.Butt, B.Raza, M.W.Ashraf and S.Begum, "Bio-inspired routing protocol for WSN-based smart grid applications in the context of Industry 4.0,"*Transactions on Emerging Telecommunications Technologies*, vol.30,no.8, pp.e3503,2019.

[26] P.S.Rao, P.K.Jana and H.Banka, "A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks,"*Wireless Network*,vol.23, no. 4, pp.2005–2020,2017.

[27] A.H. El.Bassiouny, M. Abouhawwash and H.S. Shahen, "New generalized extreme value distribution and its bivariate extension," *International Journal of Computer Applications*, vol. 173, no. 3, pp. 1-10, 2017.

[28] A.H.El.Bassiouny, M.Abouhawwash and H.S.Shahen, "Inverted exponentiated gamma and it's a bivariate extension," *International Journal of Computer Application*, vol. 3, no. 8, pp. 13-39, 2018.

[29] M. Abouhawwash and M.A. Jameel, "KKT proximity measure versus augmented achievement scalarization function," *International Journal of Computer Applications*, vol. 182, no. 24, pp. 1-7, 2018.

[30] H.S.Shahen, A.H.El.Bassiouny and M.Abouhawwash, "Bivariate exponentiated modified Weibull distribution," *Journal of Statistics Applications & Probability*, vol. 8, no. 1, pp. 27-39, 2019.

[31] M. Abouhawwash and M.A.Jameel, "Evolutionary multi-objective optimization using Benson'skush-Kuhn-tucker proximity measure," in *Proc.International Conference on Evolutionary Multi-Criterion Optimization*, East Lansing, Michigan, USA, pp. 27-38, 2019.

[32] M.Abdel Basset, R. Mohamed, M. Elhoseny, M. Abouhawwash, Y. Nam *et al*., "Efficient MCDM model for evaluating the performance of

commercial banks: A case study," *Computers, Materials & Continua*, vol. 67, no. 3, pp. 2729-2746, 2021.

[33] B.Gomathi, S.Balaji, V.K.Kumar, M.Abouhawwash, S. Aljahdali *et al*., "Multi-objective optimization of energy aware virtual machine placement in the cloud data center," *Intelligent Automation & Soft Computing*, vol.33, no.3, pp. 1771-1785, 2022.

[34] M.Kumar, K.Venkatachalam, M. Masud and M. Abouhawwash, "Novel dynamic scaling algorithm for energy-efficient cloud computing," *Intelligent Automation & Soft Computing*, vol.33, no.3, pp. 1547-1559, 2022.