Classifying Parent-Child Play Interactions Using 3D Pose Estimation and Body Orientation Analysis in Non-Face-Visible Environments

Hyeonwoo Kim

Department of Computer Science and Engineering, Soonchunhyang University, South Korea

Recognizing and interpreting parent-child interactions during play is essential for assessing children's developmental and social behaviors [1]. Although many existing computational approaches rely on facial expressions, gaze direction, or vocal signals to analyze interaction patterns, these modalities are often missing or unreliable in real-world surveillance footage. In particular, videos captured by ceiling-mounted CCTV cameras frequently depict subjects from behind, at a distance, or in low resolution, where faces are either occluded or outside the field of view [2]. To address these constraints, we propose a method that identifies parent-child play interactions using 3D human pose estimation and body orientation analysis from monocular RGB video. We extract full-body skeletal joint coordinates and body orientation vectors of both parent and child using state-of-the-art pose estimation models [3]. From this data, we calculate not only the absolute poses but also the frame-wise distances between corresponding joints across the two individuals, such as the spatial distance between each of their wrists or shoulders. These relational features, along with the raw poses and body orientations, are concatenated and used as input to a Long Short-Term Memory (LSTM)-based model that captures temporal dynamics and evolving interaction patterns [4]. We validate this approach on a dataset of naturalistic parent-child play scenarios recorded under CCTV-like conditions. Each sequence is annotated with one of several interaction types, including toy play, ball passing, and no interaction. The proposed model accurately classifies these interaction types, even when facial visibility is fully obstructed, demonstrating the effectiveness of combining absolute and relational pose information over time. This study presents a privacy-aware and context-adaptable approach to social interaction recognition, expanding the practical usability of vision-based behavior analysis in environments where facial data is inaccessible. It further contributes to growing evidence that body motion and spatial orientation are robust proxies for detecting engagement and social intention in real-world settings.

References
[1] Yogman, M., Garner, A., Hutchinson, J., Hirsh-Pasek, K., Golinkoff, R. M., Baum, R., ... & COMMITTEE ON PSYCHOSOCIAL ASPECTS OF CHILD AND FAMILY HEALTH. (2018). The power of play: A pediatric role in enhancing development in young children. Pediatrics, 142(3).
[2] Kellnhofer, P., Recasens, A., Stent, S., Matusik, W., & Torralba, A. (2019). Gaze360: Physically unconstrained gaze estimation in the wild. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 6912-6921).
[3] Wu, C., Chen, Y., Luo, J., Su, C. C., Dawane, A., Hanzra, B., ... & Kuo, C. H. (2020). Mebow: Monocular estimation of body orientation in the wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3451-3461).
[4] Khobdeh, S. B., Yamaghani, M. R., & Sareshkeh, S. K. (2024). Basketball action recognition based on the combination of YOLO and a deep fuzzy LSTM network. The Journal of Supercomputing, 80(3), 3528-3553.
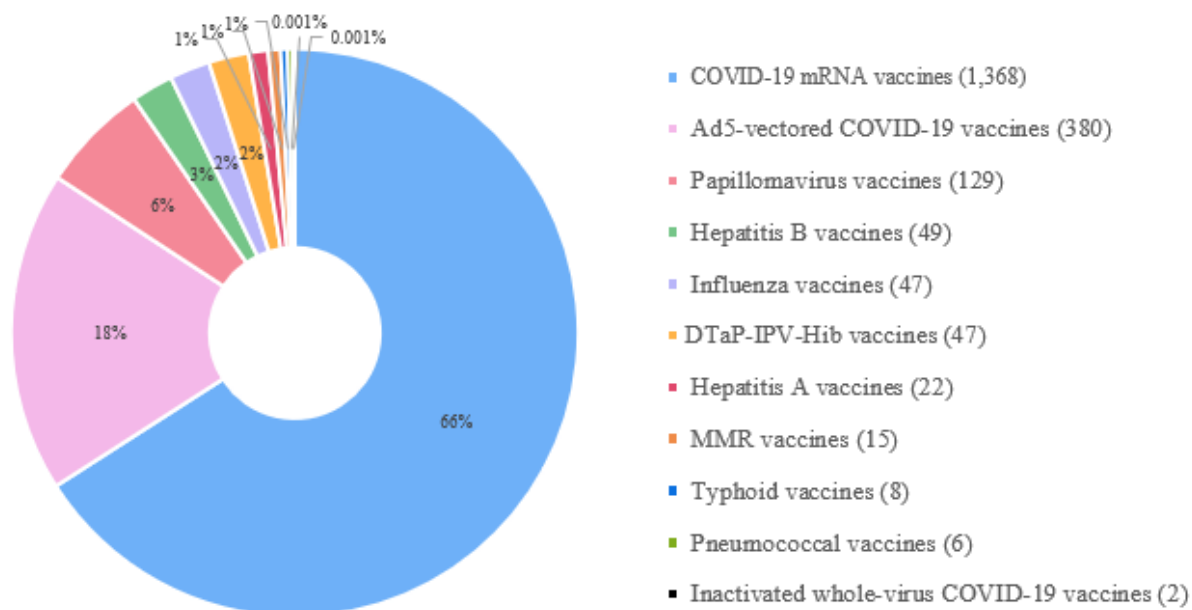
Global burden of vaccine-associated Raynaud's phenomenon: a comprehensive analysis using a pharmacovigilance database, 1968–2024

Hyunjee Kim, Yerin Hwang, Selin Woo, Dong Keon Yon

Center for Digital Health, Medical Science Research Institute, Kyung Hee University Medical Center, Kyung Hee University College of Medicine, Seoul, South Korea
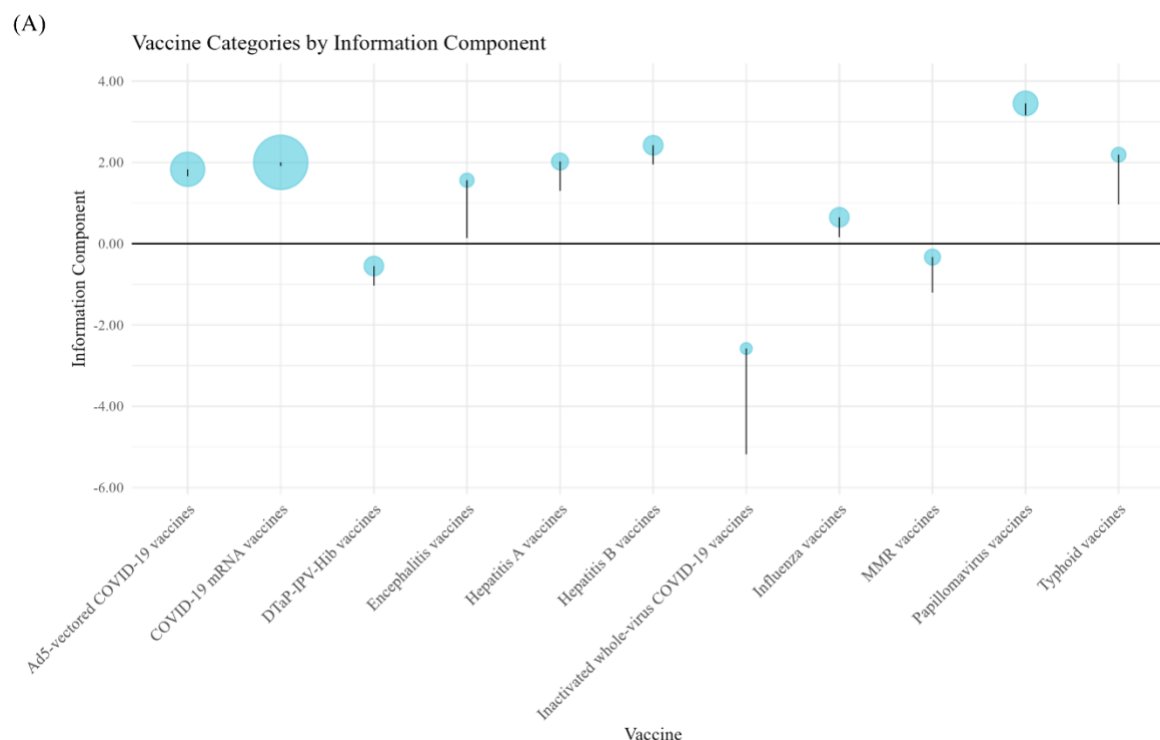
Although multiple cases of vaccine-associated secondary Raynaud's phenomenon have been reported, there is no comprehensive and systemic analysis highlighting the association between vaccines and secondary Raynaud's phenomenon.[1, 2] Therefore, this study aimed to find out the potential association between various vaccines and secondary Raynaud's phenomenon, by utilizing data from a global pharmacovigilance database, which comprise over 140 million safety reports from more than 170 countries from 1968 to 2024.[3] Reports of secondary Raynaud's phenomenon were identified using MedDRA version 26.0 terminology, and only reports in which the vaccine was classified as "suspected" were included in the analysis.[4] The 11 vaccine categories examined were as follows: (1) diphtheria, tetanus, acellular pertussis, polio, and *Haemophilus influenzae* type b (DTaP-IPV-Hib); (2) typhoid; (3) encephalitis; (4) influenza; (5) hepatitis A; (6) hepatitis B; (7) measles, mumps, and rubella (MMR); (8) papillomavirus (HPV); (9) COVID-19 mRNA; (10) adenovirus type 5 (Ad5)-vectored COVID-19; and (11) inactivated whole-virus COVID-19 vaccines. Two established disproportionality measures were employed: the reporting odds ratio (ROR) with 95% confidence intervals (CI), and the information component (IC) along with its lower credibility limit ($IC_{0.25}$). The value of $IC_{0.25}$ is >0.00 or the lower end of the ROR 95% CI is >1.00 means that it is statistically significant.[5] A total of 2,122 adverse events related to vaccine-associated Raynaud's phenomenon were identified and included in the analysis. COVID-19 mRNA vaccines accounted for the largest proportion of the reports (65.99%), followed by Ad5-vectored COVID-19 (18.33%) and papillomavirus vaccines (6.22%) (**Figure 1**). However, HPV vaccine exhibited the highest association (ROR: 11.49 [95% CI, 9.66–13.67]; IC: 3.45 [$IC_{0.25}$, 3.16]), followed by typhoid (ROR: 5.86 [95% CI, 2.93–11.72] ; IC: 2.19 [$IC_{0.25}$, 0.97]),), hepatitis B (ROR: 5.63 [95% CI, 4.25–7.45]; IC: 2.42 [$IC_{0.25}$, 1.95]), COVID-19 mRNA (ROR: 5.00 [95% CI, 4.70–5.31]; IC: 2.00 [$IC_{0.25}$, 1.91]), and hepatitis A (ROR: 4.35 [95% CI, 2.87–6.62]; IC: 2.02 [$IC_{0.25}$, 1.30]) (**Figure 2**). The association was higher in females (ROR: 3.74 [95% CI, 3.54–3.95]; IC: 1.67 [$IC_{0.25}$, 1.59]) compared to males (ROR: 3.44 [95% CI, 3.12–3.78]; IC, 1.57 [$IC_{0.25}$, 1.43]) and increased with advancing age: 0–11 years (IC, 0.03 [$IC_{0.25}$, –0.56]); 12–17 years (IC, 1.54 [$IC_{0.25}$, 1.25]); 18–44 years (IC, 1.64 [$IC_{0.25}$, 1.52]); 45–64 years (IC, 2.00 [$IC_{0.25}$, 1.87]); ≥65 years (IC, 2.12 [$IC_{0.25}$, 1.91]). This comprehensive analysis suggests a potential association between Raynaud's phenomenon and various vaccines, particularly HPV and COVID-19 mRNA vaccines, based on long-term global pharmacovigilance data from 1968 to 2024. Moreover, the significant associations in females and with increasing age groups underscore the need for targeted post-vaccination monitoring. While these findings highlight important safety signals, causality cannot be established due to the inherent limitation of the VigiBase. Therefore, further studies are warranted to better understand and address the burden of secondary Raynaud's phenomenon due to vaccinations.
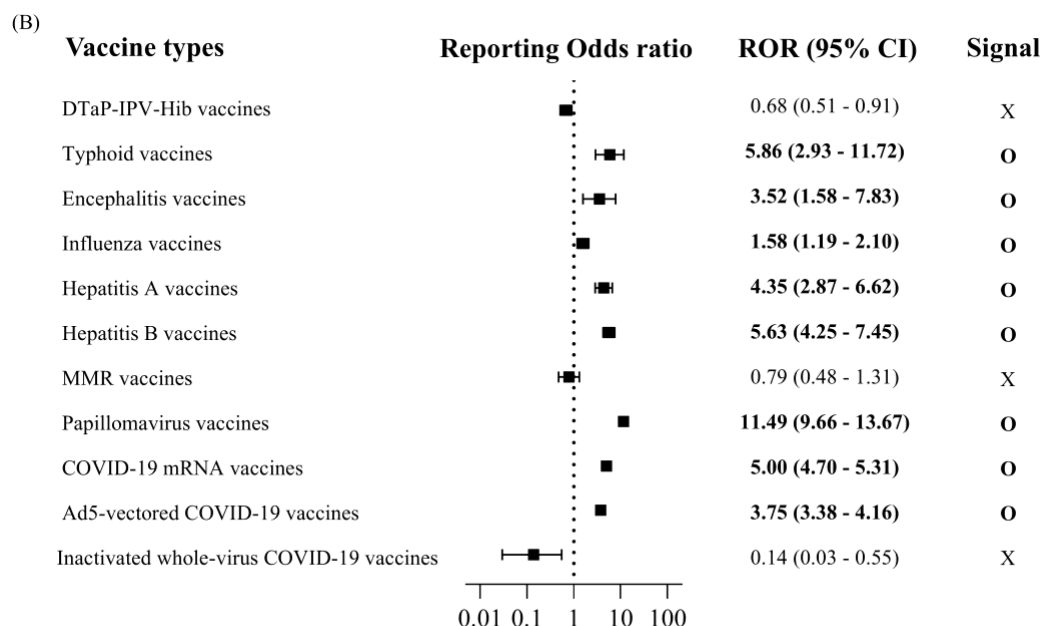
Figure 1. Proportion distribution of reports of Raynaud's phenomenon adverse events with different vaccines

Abbreviations: DTaP-IPV-Hib, diphtheria, tetanus toxoids, pertussis, polio, and *Hemophilus influenza* type b; MMR, measles, mumps, and rubella.

Figure 2. Analysis of subgroups based on $IC_{0.25}$ (A) and ROR (B) values in vaccine-associated Raynaud's phenomenon adverse events disproportionality.

(A)

(B)

Vaccine types	Reporting Odds ratio	ROR (95% CI)	Signal
DTaP-IPV-Hib vaccines		0.68 (0.51 - 0.91)	X
Typhoid vaccines		**5.86 (2.93 - 11.72)**	O
Encephalitis vaccines		**3.52 (1.58 - 7.83)**	O
Influenza vaccines		**1.58 (1.19 - 2.10)**	O
Hepatitis A vaccines		**4.35 (2.87 - 6.62)**	O
Hepatitis B vaccines		**5.63 (4.25 - 7.45)**	O
MMR vaccines		0.79 (0.48 - 1.31)	X
Papillomavirus vaccines		**11.49 (9.66 - 13.67)**	O
COVID-19 mRNA vaccines		**5.00 (4.70 - 5.31)**	O
Ad5-vectored COVID-19 vaccines		**3.75 (3.38 - 4.16)**	O
Inactivated whole-virus COVID-19 vaccines		0.14 (0.03 - 0.55)	X

0.01 0.1 1 10 100

Abbreviations: DTaP-IPV-Hib, diphtheria, tetanus toxoids, pertussis, polio, and *Hemophilus influenza* type b; MMR, measles, mumps, and rubella.

References

[1]   Curtiss P, Svigos K, Schwager Z, Lo Sicco K, Franks AG, Jr. (2024) Part II: The treatment of primary and secondary Raynaud's phenomenon. J Am Acad Dermatol 90 (2): 237-248 DOI 10.1016/j.jaad.2022.05.067

[2]   Curtiss P, Svigos K, Schwager Z, Lo Sicco K, Franks AG, Jr. (2024) Part I: Epidemiology, pathophysiology, and clinical considerations of primary and secondary Raynaud's phenomenon. J Am Acad Dermatol 90 (2): 223-234 DOI 10.1016/j.jaad.2022.06.1199

[3]   Cho J, Jo H, Park J, Oh J, Kim TH, Lee K, Lee H, Jeong J, Lee S, Miligkos M, Papadopoulos NG, Yon DK (2025) TOP10-SCAR: A Global Pharmacovigilance Study on Medications Most Frequently Related to Severe Cutaneous Adverse Reactions. Allergy   DOI 10.1111/all.16544

[4]   Jeong YD, Lee K, Park J, Lee J, Kang J, Yeo SG, Smith L, Lee H, Yon DK (2025) Global burden of vaccine-associated angioedema and their related vaccines, 1967-2023: Findings from the global pharmacovigilance database. Allergy 80 (2): 590-594 DOI 10.1111/all.16304

[5]   Kyung S, Woo S, Kim M, Lee H, Kang J, Rahmati M, Yon DK (2024) Global burden of vaccine-associated alopecia, 1979-2023: a comprehensive analysis of the international pharmacovigilance database. Br J Dermatol 190 (5): 764-767 DOI 10.1093/bjd/ljae055

Optimal 3D Node Localization in Dynamic Wireless Sensor Networks

Hyung Ju Park, Yunyoung Nam, Byeong-Gwon Kang

Department of ICT convergence, Soonchunhyang University, Asan 31538, South Korea

Contact: Yunyoung Nam, Byeong-Gwon Kang. Email: ynam@sch.ac.kr, bgkang@sch.ac.kr

Abstract— **Localization is critical for many Wireless Sensor Network (WSN) applications, particularly in three-dimensional (3D) environments where two-dimensional (2D) models are inadequate due to significant altitude variations. This paper proposes a novel 3D localization method using a single anchor node to determine the positions of mobile nodes in a dynamic WSN. The nodes, deployed in an anisotropic environment with a Degree of Irregularity (DOI) of 0.01, move across middle and lower layers, while a single anchor node is positioned at the top layer. We introduce the Adaptive Plant Propagation Algorithm (APPA), a soft computing technique, to optimize the localization of these heterogeneous mobile nodes. Simulation results demonstrate that APPA outperforms other meta-heuristic algorithms, such as PSO, HPSO, BBO, GWO, and FA, in terms of localization error, computational efficiency, and the number of successfully located nodes.**

Keywords: Wireless Sensor Networks, 3D Localization, Single Anchor Node, Adaptive Plant Propagation Algorithm, Meta-Heuristic Optimization

I. INTRODUCTION

Wireless Sensor Networks (WSNs) consist of numerous low-power sensor nodes (SNs) deployed to monitor physical phenomena, such as temperature, humidity, or location [1]. Accurate localization is essential, as sensor data without positional context is often meaningless [2]. While Global Positioning System (GPS) is widely used, it fails in environments with obstructed line-of-sight (LOS), such as dense forests, mountains, or indoor settings [2]. To address this, WSNs employ anchor nodes with known positions to localize unknown nodes using range-based or range-free algorithms [3-4]. Range-based methods, such as Received Signal Strength Indicator (RSSI), Angle of Arrival (AoA), or Time of Arrival (ToA), offer precision but require additional hardware. Range-free methods, like Distance Vector-Hop or Ad Hoc Positioning, reduce infrastructure needs but are less accurate.

In dynamic 3D WSNs, where nodes move and altitude varies, localization is particularly challenging. Most existing techniques focus on 2D scenarios, leaving 3D localization underexplored. This paper proposes a novel localization approach using a single anchor node, augmented by six virtual anchor nodes positioned at 60-degree intervals, to localize mobile nodes in a 3D anisotropic environment. We employ the Adaptive Plant Propagation Algorithm (APPA), a meta-heuristic inspired by plant propagation, to optimize node positions. APPA leverages a single anchor's RSSI measurements and virtual anchor configurations to achieve high accuracy with minimal infrastructure.

II. RELATED WORK

Localization in WSNs has been extensively studied. Liu et al. [5-6] explored Mobile Ad Hoc and Opportunistic Networks, but their work is less relevant to WSNs with multiple sensor nodes. Recent research primarily addresses 2D localization, which is inadequate for 3D environments with complex topologies, such as mountainous or aerial deployments.

Chu et al. [7] introduced the Multi-Group Quantum-Behavior Symbiotic Organism Search (MQSOS), outperforming algorithms like PSO and QUATRE on the CEC2013 benchmark. Liu et al. [8] proposed distributed localization using inter-node distance data, achieving high accuracy and energy efficiency. Kotwal et al. [9] used RSSI-based distance estimates and PSO to refine node positions within feasibility regions, requiring fewer anchor nodes. Low et al. [10] applied PSO to localize emitters using RSSI from four anchor nodes, demonstrating real-time applicability. Wang et al. [11] combined Bacterial Foraging Algorithm (BFA) and Glow-worm Swarm Optimization (GSO) for improved convergence. Graefenstein et al. [12] and Sumathi et al. [13] utilized RSSI-based trilateration and least-squares methods, respectively, for single-anchor localization. Guo et al. [14] proposed a mobile beacon-based Perpendicular Intersection (PI) method, while Shi et al. [15] used ultra-wideband (UWB) signals for localization. Other approaches include DV-Hop [16], quasi-Newton optimization [17], irregular RSSI models [18], parametric loop-division [19], and swarm intelligence [20-25]. PSO-based methods [26-30] have shown promise for

mobile node localization but often require multiple anchor nodes.

This paper advances the field by using a single anchor node with virtual anchors and APPA to achieve efficient 3D localization in dynamic WSNs, addressing the limitations of multi-anchor and static-node approaches.

III. ADAPTIVE PLANT PROPAGATION ALGORITHM (APPA)

The Adaptive Plant Propagation Algorithm (APPA) is a population-based meta-heuristic inspired by the propagation of strawberry plants. Each solution (plant) in the search space is evaluated using an objective function, and solutions propagate via runners to explore or exploit the search space. Plants in poor locations send few, long runners (exploration), while those in favorable locations send many, short runners (exploitation).

Let the search space have dimension D, population size NP, maximum generations G_{max}, and maximum runners per plant N_{max}. The fitness of a solution x_i is mapped to [0,1] using the sigmoid function:

$$f(x_i) = \frac{1}{1+e^{-g(x_i)}} \tag{1}$$

where $g(x_i)$ is the objective function value. The number of runners for solution x_i is:

$$N_r(x_i) = [f(x_i) \cdot N_{max} \cdot \text{rand}()] \tag{2}$$

where $\text{rand}() \in [0,1]$, ensuring at least one runner. The runner distance is inversely proportional to fitness:

$$d_j = \text{rand}() \cdot (U_j - L_j) \cdot (1 - f(x_i)) \tag{3}$$

where L_j and U_j are the lower and upper bounds of the j-th dimension. The new position is updated as:

$$x_{i,\text{new},j} = x_{i,j} + d_j \cdot \text{rand}() \tag{4}$$

To prevent solutions from exiting the search space, APPA adaptively adjusts positions to stay within bounds. To avoid local minima, solutions stagnating for a threshold number of generations are replaced with new random solutions. This adaptive mechanism enhances APPA's robustness for 3D localization.

IV. SINGLE ANCHOR NODE LOCALIZATION

In the proposed 3D localization scheme, a single anchor node with known coordinates (x_a, y_a, z_a) localizes mobile nodes in a cubic 3D environment divided into three layers. The anchor resides at the top layer, while 40 mobile nodes per layer occupy the middle and lower layers. Six virtual anchor nodes are placed at 60-degree intervals around the anchor at the same distance, forming an umbrella projection. For each target node, three virtual anchors and the real anchor are selected based on directional information to compute 3D coordinates.

The distance between the anchor and a target node at (x_t, y_t, z_t) is estimated using RSSI:

$$d = \sqrt{(x_t - x_a)^2 + (y_t - y_a)^2 + (z_t - z_a)^2} \tag{5}$$

The centroid of the selected anchors (real and virtual) is calculated as:

$$(x_c, y_c, z_c) = \left(\frac{\sum_{i=1}^{4} x_i}{4}, \frac{\sum_{i=1}^{4} y_i}{4}, \frac{\sum_{i=1}^{4} z_i}{4}\right) \tag{6}$$

APPA uses this centroid as the initial guess to optimize the target node's position, minimizing the objective function:

$$f(x_t, y_t, z_t) = \sum_{i=1}^{4} \left(\sqrt{(x_t - x_i)^2 + (y_t - y_i)^2 + (z_t - z_i)^2} - d_i\right)^2 \tag{7}$$

where d_i is the estimated distance to the i-th anchor. This approach ensures accurate localization with minimal infrastructure.

V. SIMULATION RESULTS AND DISCUSSION

The proposed APPA-based localization method was evaluated in a 3D cubic environment with 80 mobile nodes (40 per lower layer) and a single anchor node at the top layer. Six virtual anchors form an umbrella projection, with distances estimated via RSSI (DOI = 0.01). Table 1 lists the parameter settings for APPA and competing algorithms (PSO, HPSO, BBO, GWO, FA):

Table 1: Parameter settings

Algorithm	Parameters
PSO	NP=30; D=3; Gmax=100; c_1, c_2, c_3=1.494; w=0.729
HPSO	NP=30; D=3; Gmax=100; c_1, c_2, c_3=1.494; η=0.1; w=0.729
BBO	NP=30; D=3; Gmax=100; p_m =0.05
FA	NP=30;D=3; Gmax=100; $\alpha = 0.5$; $\gamma = 1.0$
GWO	NP=20;D=3; Gmax=100; a= 2.0
APPA	NP=30;D=3; Gmax=100; $n_{max} = 5$

The average localization error is computed as:

$$E_{\text{avg}} = \frac{1}{N}\sum_{i=1}^{N} \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (z_i - \hat{z}_i)^2} \tag{8}$$

where (x_i, y_i, z_i) and $(\hat{x}_i, \hat{y}_i, \hat{z}_i)$ are the actual and estimated positions, respectively. Table 2 compares the performance across five movement scenarios:

Table 2: Comparison of meta-heuristic algorithms

Algorithm	Movements	Max Error	Min Error	Avg Error	Located Targets
PSO	1–5	3.9358–5.3379	0.0367–0.0831	0.9267–0.9958	80
HPSO	1–5	3.1204–5.2376	0.0230–0.1044	0.4032–0.6742	80
BBO	1–5	5.3500–5.9014	0.1528–0.3318	1.1585–1.2818	80
GWO	1–5	3.1101–4.9834	0.0430–0.0944	0.3932–0.6442	80
FA	1–5	6.1101–6.9036	0.1627–0.3412	2.2013–2.5123	80
APPA	1–5	2.8232–4.3076	0.0353–0.0674	0.2252–0.4474	80

APPA consistently achieves the lowest average localization error (0.2252–0.4474) and faster convergence, outperforming PSO, HPSO, BBO, GWO, and FA. The single-anchor approach with virtual anchors mitigates LOS issues, and APPA's adaptive mechanism enhances robustness in dynamic scenarios. Future work could explore hybridizing APPA with other algorithms to further reduce errors.

VI. Conclusions

This paper presents a novel 3D localization method for dynamic WSNs using a single anchor node and six virtual anchors forming an umbrella projection. The APPA algorithm optimizes node positions by minimizing RSSI-based localization errors. Simulation results demonstrate that APPA outperforms competing meta-heuristic algorithms in accuracy and convergence speed. The proposed method is applicable to diverse scenarios, including logistics, underwater monitoring, and remote event localization. Future enhancements could involve hybrid optimization techniques to further improve accuracy.

Acknowledgment

References

[1] R. Kulkarni, G. Venayagamoorthy and M. Cheng, "Bio-inspired node localization in wireless sensor networks," in 2009 IEEE International Conference on Systems, Man and Cybernetics, vol. 22, pp. 205-210, 2009.

[2] D. Lavanya and S. K. Udgata, "Swarm intelligence based localization in wireless sensor networks," International Workshop on Multi-Disciplinary Trends in Artificial Intelligence, Springer, Berlin, Heidelberg, pp. 317-328, 2011.

[3] P. Singh, B. Tripathi and N. Singh, "Node localization in wireless sensor networks," International journal of computer science and information technologies, vol. 2, no. 6, pp. 2568-2572, 2011.

[4] A. Boukerche, H. Oliveira, E. Nakamura and A. Loureiro, "Localization systems for wireless sensor networks," IEEE wireless Communications, vol. 14, no. 6. pp, 6-12, 2006.

[5] Y. Liu, Z. Yang, T. Ning and W. Hongyi, "Efficient quality-of-service (QoS) support in mobile opportunistic networks," IEEE Transactions on Vehicular Technology, vol. 63, no. 9, pp. 4574-4584, 2014.

[6] Y. Liu, Z. Yang, T. Ning and W. Hongyi, "Efficient data query in intermittently-connected mobile ad hoc social networks," IEEE Transactions on Parallel and Distributed Systems, vol. 26, no. 5, pp. 1301-1312, 2014.

[7] S. Chu, Z. Du and J. Pan, "Symbiotic Organism Search Algorithm with Multi-Group Quantum-Behavior Communication Scheme Applied in Wireless Sensor Networks," MDPI, vol. 10, no. 3, pp. 930-952, 2020.

[8] D. Liu, S. Guo, W. Chen and F. Wang, "History based multi-node collaborative localisation in mobile wireless ad hoc networks," International Journal of Ad Hoc and Ubiquitous Computing, vol. 30, no. 2, pp. 59-72, 2019.

[9] S. Kotwal, S. Gill and K. Saini, "Development of range free three dimensional localisation in wireless sensor networks," International Journal of Sensor Networks, vol. 31, no. 1, pp. 52-63, 2019.

[10] K. Low, H. Nguyen and H. Guo, "Optimization of sensor node locations in a wireless sensor network," in Fourth international conference on natural computation, IEEE, vol. 5, pp. 286-290, 2008.

[11] Y. Wang, P. Wang, J. Zhang, X. Cai, W. Li et al., "A novel DV-Hop method based on coupling algorithm used for wireless sensor network localisation," International Journal of Wireless and Mobile Computing, vol. 16, no. 2, pp. 128-137, 2019.

[12] J. Graefenstein, A. Albert, P. Biber and A. Schilling, "Wireless node localization based on RSSI using a rotating antenna on a mobile robot," in 2009 6th Workshop on Positioning, Navigation and Communication, IEEE, pp. 253-259, 2009.

[13] R. Sumathi and R. Srinivasan, "RSS-based location estimation in mobility assisted wireless sensor networks," in Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems, vol. 2, no. 3, pp. 848-852, 2011.

[14] Z. Guo, Y. Guo, F. Hong, Z. Jin, Y. He et al., "Perpendicular intersection: locating wireless sensors with mobile beacon," IEEE Transactions on Vehicular Technology, vol. 59, no. 7, pp. 3501-3509, 2010.

[15] Q. Shi, H. Huo, T. Fang and D. Li, "A 3d node localization scheme for wireless sensor networks," IEICE Electron Exp, vol. 6, no. 3, pp. 67–72, 2009.

[16] L. Wang, J. Zhang and D. Cao , "A new 3-dimensional dv-hop localization algorithm," Journal of Computing Information Systems, vol. 8, no.6, pp. 2463–75,2012.

[17] Y. Xu, Y. Zhuang and J. Gu, "An improved 3D localization algorithm for the wireless sensor network", International Journal of Distributed Sensor Networks, vol. 11, no. 6, pp. 1-13, 2015.

[18] J. Li, X. Zhong and I. Lu, "Three-dimensional node localization algorithm for WSN based on differential RSS irregular transmission model," Journal of Communications vol. 9, no. 5, pp. 391-397, 2014.

[19] T. Ahmad, X. Li and B. Seet, "Parametric loop division for 3D localization in wireless sensor networks," Sensors, MDPI vol. 17, no. 7 pp. 1697-1729, 2017.

[20] A. Gopakumar and L. Jacob, "Localization in wireless sensor networks using particle swarm optimization," in IET Conference on Wireless, Mobile and Multimedia Networks, pp. 227-230. 2008.

[21] P. Chuang and C. Wu, "An effective pso-based node localization scheme for wireless sensor networks," Ninth international conference on parallel

and distributed computing, applications and technologies, in PDCAT, IEEE, pp. 187–194, 2008.

[22] J. Kennedy, "Bare bones particle swarms," Swarm Intelligence Symposium, SIS'03. IEEE, pp. 80-87, 2003.

[23] R. Kulkarni and G. Venayagamoorthy, "Particle swarm optimization in wireless-sensor networks A brief survey," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 41, no.2, pp. 262-267, 2011.

[24] A. Kumar, A. Khosla, J. Saini and S. Singh, "Meta-heuristic range-based node localization algorithm for wireless sensor networks," in International conference on localization and GNSS (ICL-GNSS), IEEE, pp. 1-7, 2012.

[25] S. Arora and S. Singh, "Node localization in wireless sensor networks using butterfly optimization algorithm," Arabian Journal of Science and Engineering, vol. 42, no. 8, pp. 1–11, 2017.

[26] P. Singh, A. Khosla, A. Kumar and M. Khosla, "A Novel Approach for Localization of Moving Target Nodes in Wireless Sensor Networks," International Journal of Grid and Distributed Computing, vol. 10, no. 10, pp. 33-44, 2017.

[27] P. Singh, A. Khosla, A. Kumar and M. Khosla, "Computational intelligence based localization of moving target nodes using single anchor node in wireless sensor networks." Telecommunication Systems, Springer vol. 69, no. 3, pp. 397-411, 2018.

[28] P. Singh, A. Khosla, A. Kumar and M. Khosla, "3D localization of moving target nodes using single anchor node in anisotropic wireless sensor networks," AEU-International Journal of Electronics and Communications, vol. 82, no. 1, pp. 543-552, 2017.

[29] P. Singh, A. Khosla, A. Kumar and M. Khosla, "Optimized localization of target nodes using single mobile anchor node in wireless sensor network," AEU-International Journal of Electronics and Communications, vol. 91, no. 3, pp.55-65, 2018.

[30] P. Singh, A. Khosla, A. Kumar and M. Khosla, "A novel approach for localization of moving target nodes in wireless sensor networks," International Journal of Grid and Distributed Computing, vol. 10, no. 10, pp. 33-44, 2017.

Deep Learning for Predicting Hospitalization of Abdominal Symptom Patients in the Emergency Department using combining initial abdominal radiography and early clinical information

Ahsan Aziz[1], Yeo Eun Han[2], Yunyoung Nam[3] and Yongwon Cho[3]

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

[2]Department of Radiology, Korea University Anam Hospital, 73 Goryeodae-ro, Seongbuk-gu, Seoul, 02841, Republic of Korea

[3]Department of Computer Science and Engineering, Soonchunhyang University, Asan, 31538, Korea

Contact: dragon1won@sch.ac.kr

Abstract:

Abdominal pain has a wide variety of causes and treatments. Some conditions, such as gallstones or appendicitis, may require surgery. Others, such as ulcers or infections, may be relieved with medicine. Abdominal imaging is used to detect and diagnose conditions such as tumors, infections, gastrointestinal disorders, kidney stones, and abdominal trauma. It helps monitor disease progression and guide surgical interventions. In [1], they evaluate major abdominal complications (MACs) after penetrating abdominal trauma (PAT) using 4,473 trauma patients, which includes 2,326 firearm injuries and 2,147 stab injuries, with main variables included demographics, injury severity, treatment approach, and complications. In [2], they analyze unnecessary admissions for abdominal pain under an Acute General Surgical Unit (AGSU) using 1,587 patients admitted for abdominal pain with key variables as length of stay (LOS), blood markers (CRP, WCC), and imaging results. We have a raw abdominal X-ray dataset from which we have extracted two classes, admission and discharge, while applying the pre-processing and augmentation techniques on the data. Along with the image, the clinical text dataset is also available. Our methodology consists of three scenarios: firstly, we will only use ML classifiers secondly, we will use CNN only for the classification purpose separately and lastly, we will be using CNN to extract image features and LLM for extraction of text embeddings, then we'll have a fusion layer where the fusion process will be performed and combines both features for final classification. for evaluation we will compare the performances of all three approaches and analyze the contribution of the text features in improving the accuracy.

Keywords: Abdominal Pain, Deep Learning, Machine Learning, Multimodal, Fusion.

Introduction:

Abdominal pain is a prevalent symptom experienced by individuals worldwide, with various causes and significant implications for healthcare systems. Approximately 51.9% of the global population reported experiencing abdominal pain in the last three months according to [4] and in the United States, abdominal pain is the leading gastrointestinal symptom prompting ambulatory visits. Some of the common causes of abdominal pain are gastrointestinal disorders (Conditions such as irritable bowel syndrome (IBS), gastritis, peptic ulcers, and diverticular disease are common culprits, [5]infections (Bacterial infections like Helicobacter pylori can lead to gastritis and ulcers), reproductive system issues (In women, conditions like endometriosis, fibroids, and pelvic inflammatory disease can cause significant abdominal discomfort) and functional disorders (Functional abdominal pain disorders (FAPDs) are prevalent, especially among children, with varying global prevalence rates). Irritable Bowel Syndrome, or IBS, is a communal ailment that affects an estimated 10-15 percent of the world population. While it can cause minor to reasonable discomfort in some people, others may experience more stern side effects like abdominal pain and bloating and it is a communal yet complex disorder of the digestive system. It is assessed that it affects up to 15% of the population, with women being more likely to suffer from it. Although uncomfortable and disruptive, IBS does not pose any long-term damage to the intestines or other organs, making it a chronic yet manageable condition.

Related Work:

In [1] the main objective was to evaluate the burden of major abdominal complications (MACs) after penetrating abdominal trauma (PAT). The authors have contributed by identifying predictors of MACs after PAT and highlighting differences in MACs

between firearm injuries (FIs) and stab injuries (SIs). their proposed methodology was that data from the 2012-2015 National Readmission Database and patients with penetrating injuries undergoing exploratory laparotomy were tracked for 6 months post-discharge then statistical analysis including regression analysis. the dataset authors used was 4,473 trauma patients (2,326 FIs; 2,147 SIs) and variables included demographics, injury severity, treatment approach, and complications. they have concluded that 22% of patients developed MACs postoperatively, fIs had a higher risk of MACs compared to SIs, and predictors of MACs: Damage control laparotomy (DCL), bowel perforation, biliary-pancreatic injury, hepatic injury, and blood transfusion. They have achieved the results as IAA (intra-abdominal abscess) in 19%, superficial SSI in 7%, fascial dehiscence in 4%, and 6-month mortality was 7%, higher for firearm injuries.
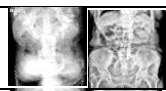
Firearm injuries resulted in greater organ damage and higher risk of MACs compared to stab injuries. The limitations Only include hospital readmissions within the same state and limited follow-up data beyond 6 months. In [2] the main objective was to analyze unnecessary admissions for abdominal pain under an Acute General Surgical Unit (AGSU). they proposed a retrospective cohort study over two years at a single hospital in Victoria, Australia, with data extracted from electronic medical records, and statistical analysis of diagnostic markers, imaging, length of stay, and pain outcomes. their main contribution was to identify inefficiencies in AGSU and factors that contribute to unnecessary hospital stays and suggest a more efficient model for managing non-surgical cases. The dataset consisted of 1,587 patients admitted for abdominal pain (70.3% had non-surgical diagnoses) along key variables such as length of stay (LOS), blood markers (CRP, WCC), and imaging results. By achieving the results as Non-surgical patients had a shorter LOS (25.3 hours) than surgical patients (44.2 hours) and CRP and imaging findings were significant predictors of surgical cases. And concluded that many admissions were unnecessary and could be managed as outpatients, patients with normal CRP and imaging results were likely to be non-surgical, and suggested incorporating CRP and imaging findings in admission protocols. In [3] the author's objective was to investigate discharge outcomes among elderly patients undergoing emergency abdominal surgery and contributed by examining factors affecting discharge destinations, length of stay, and nursing home admissions and exploring the impact of nursing home capacity on postoperative hospital stay. Their proposed approach was the registry study using

National Quality Assurance Improvement System (NQAIS) data (2014-2018) in Ireland and statistical analysis, including regression modeling. On a dataset of 4,951 emergency surgery cases (patients aged 65+) and variables such as demographics, Charlson Comorbidity Index, admission source, and discharge destination they achieved higher mortality in older patients and those transferred from other hospitals, 20% of patients were discharged to nursing homes and patients admitted from home and later discharged to nursing homes had significantly longer hospital stays. And concluded that 12.7% of elderly patients died in the hospital, and patients discharged to nursing homes had significantly longer hospital stays, and also, nursing home capacity was negatively associated with post-op length of stay.

Methods and materials:

The dataset employed in this study comprises abdominal X-ray images sourced from hospital records, categorized into two clinical groups: Admission and Discharge. Initially, the dataset contained 2,661 images, with 821 images in the admission class and 1,840 in the Discharge class, reflecting a natural clinical imbalance. To address this issue and enhance the robustness of the deep learning models, data augmentation techniques, including rotation, flipping, scaling, and transposing, were applied specifically to the underrepresented Admission class. As a result, the number of Admission images increased to 1,800, improving class balance. Table 1 presents a comprehensive overview of the dataset, including the original image count, the number of images after augmentation, and representative sample images from each class, illustrating the visual diversity and complexity tackled in this study. To ensure balanced representation across all data subsets, the dataset was split in a stratified manner into three parts: 70% for training, 20% for validation, and 10% for testing.

Table 1 Class-wise Distribution of Abdominal X-ray Dataset Before and After Augmentation with Representative Samples

Number of Classes	Image Count (Pre-Augmentation)	Image Count (Post-Augmentation)	Sample Images
Admission	821	1800	
Discharge	1840	1840	
Total	2661	3640	

In Figure 1, we have the proposed diagram, which shows that the raw abdominal X-ray images initially undergo a specialized preprocessing step to enhance image quality and highlight important visual features by employing Probability Harmonic Mean-based Contrast Enhancement, improving contrast while preserving important structural information relevant to diagnosis. Several augmentation techniques are applied to overcome the limited size of annotated medical imaging datasets, which consist of rotation, flipping, scaling, and transpose-based transformations. These augmentations increase the diversity of the dataset and help the model generalize better. Then we utilize pretrained deep learning models (such as ResNet-50, ResNet-101, DenseNet-201, Inception V3, and DarkNet-53) for feature extraction, leveraging transfer learning to fine-tune them on the abdominal X-ray dataset, and each model extracts deep feature representations from the enhanced and augmented images. These features are being used in multiple paths, first, they are directly used for end-to-end classification, and secondly, fused with clinical data features for multimodal analysis. The raw clinical text data undergoes the preprocessing steps, which consist of the removal of corrupted or incomplete records, imputation, handling of missing values, and normalization and standardization to ensure consistency in the feature scale and format. Now, to extract semantic and contextual insights from clinical notes and metadata, we employ Bidirectional Encoder Representations from Transformers (BERT), which generates high-dimensional feature vectors that represent the underlying clinical information in a contextualized manner. The feature vectors obtained from both modalities of images (via CNN) and clinical data (via BERT) are integrated through feature-level fusion. This multimodal fusion combines visual and textual information into a unified representation, enhancing the model's ability to make informed predictions based on complementary data sources.

In the first evaluation path, predictions are made directly from the pretrained and fine-tuned deep learning models based solely on the processed X-ray images, which offers insights into the performance of the standalone DL models in a pure image-based diagnosis scenario. In the second evaluation path, the fused features from both X-ray and clinical data are input to machine learning classifiers, which include fine tree, cubic support vector machine (SVM), subspace discriminant, and subspace k-nearest neighbors (KNN) these classifiers are trained and evaluated on the combined feature space to determine the added value of clinical data in improving diagnostic accuracy.
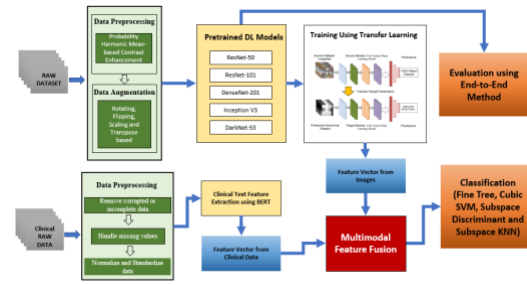

Figure 1 Proposed Diagram of our Abdominal X-ray images and Clinical data Classification

Results and discussion:

In the result section, first, we will represent the outcomes of the end-to-end deep learning approach using the abdominal X-ray images pre-augmentation and post-augmentation. Table 2 shows the performance metrics of all five deep learning models (such as ResNet-50, ResNet-101, DenseNet-201, Inception V3, and DarkNet-53), and for pre-augmentation, ResNet-50 achieved the highest accuracy of 74.25% among all five DL models, following precision rate, sensitivity, specificity, f1-score and AUC of 0.78, 0.88, 0.45, 0.82, and 0.72 respectively. And for post-augmentation, ResNet-50 again achieved the highest accuracy of 81.87% along with the highest AUC value of 0.88 among all five DL models. Figure 2 shows the AUC for ResNet-50 on pre-augmentation and post-augmentation.

Table 2 Results on End-to-End DL models using Pre-Augmentation and Post-Augmentation

| DL Model	Accuracy	Precision Rate	Sensitivity	Specificity	F1-Score	AUC
Pre-Augmentation						
ResNet-50	**74.25%**	**0.78**	**0.88**	**0.45**	**0.82**	**0.72**
ResNet-101	71.05%	0.73	0.83	0.32	0.78	0.70
Inception V3	66.73%	0.80	0.79	0.55	0.79	0.73
DenseNet-201	71.99%	0.76	0.86	0.40	0.81	0.70
DarkNet-53	71.05%	0.79	0.84	0.51	0.82	0.75
Post-Augmentation						
ResNet-50	**81.87%**	**0.81**	**0.83**	**0.81**	**0.82**	**0.88**
ResNet-101	75.69%	0.79	0.67	0.82	0.72	0.85
Inception V3	79.53%	0.75	0.84	0.71	0.79	0.86
DenseNet-201	81.04%	0.77	0.86	0.73	0.81	0.86
DarkNet-53	78.98%	0.78	0.79	0.77	0.78	0.86
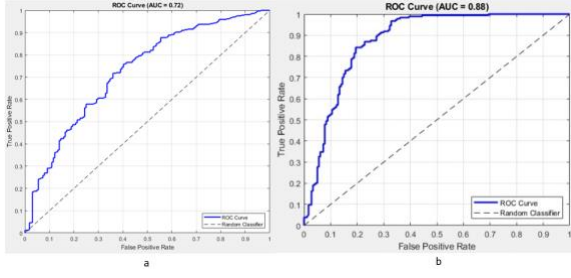
Figure 2 Pre-Augmentation AUC curve on ResNet-50 (a), Post-Augmentation AUC curve on ResNet-50 (b)

Now using ResNet-50 as our deep learning model for the extraction of X-ray image features, and also using the BERT model for the extraction of features from clinical data, and fusing them to feed them into the machine learning classifiers (such as fine tree (FT), cubic support vector machine (CSVM), subspace discriminant (SD), and subspace k-nearest neighbors (SKNN)) all the results were pre-augmentation and post-augmentation shown in Table 3 along with AUC curves in the Figure 3. The highest accuracy of 88.5% was achieved in pre-augmentation by SKNN, and also post-augmentation, SKNN achieved the highest accuracy of 74.2%. As shown in the table, the accuracy may have decreased after augmentation in our fusion approach. Still, we can see the major change across the AUC curve and the specificity of pre-post-augmentation. By having a major increase in the AUC curves and specificity values across post-Augmentation. Before Augmentation, the classifiers had higher accuracy and sensitivity, but very low specificity, indicating models were biased toward the majority class, whereas after augmentation, specificity and AUC improved for all models, which means the

models became more balanced in identifying both classes. The drop in accuracy and sensitivity is expected due to handling a more balanced dataset after augmentation. So, it is concluded that post-augmentation results are better in terms of fairness and generalization, despite a little lower accuracy. As we are targeting a real-world application where both classes are important (e.g., disease diagnosis), post-augmentation is more reliable.

Table 3 Results on the Hybrid Fusion approach using Pre-Augmentation and Post-Augmentation

| Classifiers	Accuracy	Precision Rate	Sensitivity	Specificity	F1-Score	AUC
Pre-Augmentation						
FT	81.0%	0.8894	0.8937	0.2203	0.891566	0.5261
CSVM	86.4%	0.8867	0.9688	0.0965	0.925905	0.7456
SD	81.1%	0.8870	0.8967	0.2358	0.891843	0.6486
SKNN	**88.5%**	**0.9700**	**0.9057**	**0.5455**	**0.936733**	**0.7958**
Post-Augmentation						
FT	68.4%	0.6832	0.7113	0.6559	0.696970	0.6948
CSVM	71.6%	0.6931	0.7527	0.6804	0.721649	0.8175
SD	72.6%	0.7030	0.7634	0.6907	0.731959	0.7744
SKNN	**74.2%**	**0.7525**	**0.7600**	**0.7222**	**0.756219**	**0.8572**
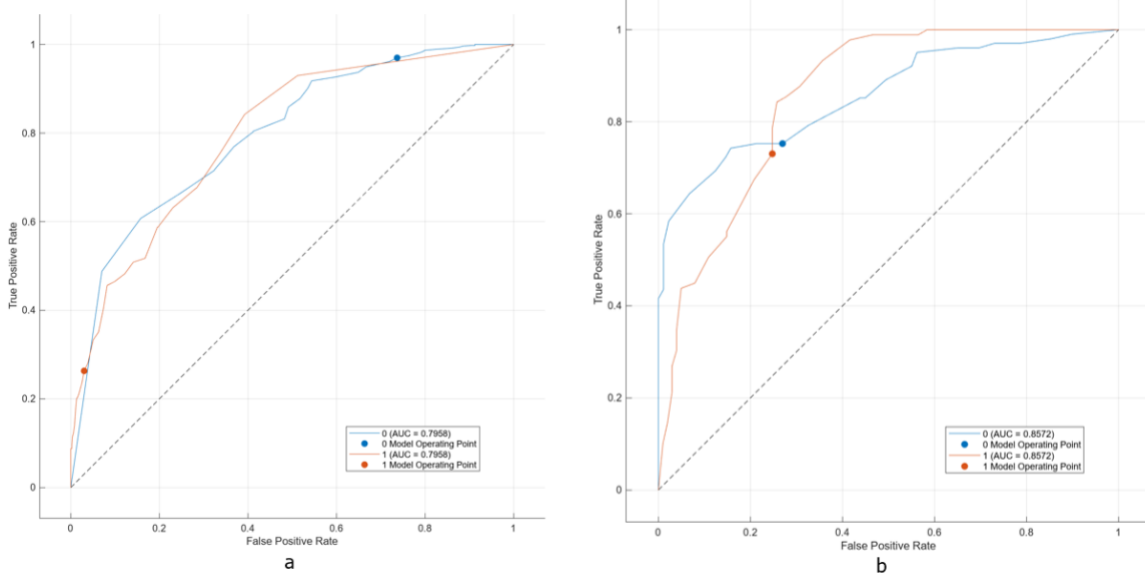
Figure 3 Pre-Augmentation AUC curve on Hybrid Fusion Approach (a), Post-Augmentation AUC curve on Hybrid Fusion Approach (b)

Conclusion:

In this study, we proposed a robust and clinically meaningful framework that integrates abdominal X-ray imaging with contextual clinical data to classify patients into Admission or Discharge categories. Our methodology follows a two-path architecture: an end-to-end deep learning model using pretrained CNNs for direct classification, and a multimodal fusion strategy where deep features from X-ray images are combined with contextual features extracted from clinical text using BERT. These fused features are then used with traditional machine learning classifiers for enhanced prediction performance. Initially, the Admission class was underrepresented, which posed a challenge for the model's learning stability and generalization. We successfully expanded the admission class through targeted augmentation techniques, improving class balance and significantly enhancing classification performance, which led to notable improvements in model accuracy, sensitivity, and overall generalization, particularly in the post-augmentation phase. Furthermore, our experimental results demonstrate that multimodal fusion consistently outperforms single-modality approaches, confirming the value of combining visual and clinical data in medical diagnostics. The dual evaluation strategy also allowed for a comparative analysis between pure deep learning and hybrid deep feature–ML pipelines, revealing the strength of each depending on the clinical scenario. In conclusion, the proposed framework not only addresses data imbalance but also demonstrates the effectiveness of leveraging both imaging and clinical modalities for improved decision-making in healthcare. This work lays a strong foundation for future extensions such as real-time clinical decision support systems and explainable AI in radiological workflows.

References:

[1] Hanna, Kamil, Samer Asmar, Michael Ditillo, Mohamad Chehab, Muhammad Khurrum, Letitia Bible, Molly Douglas, and Bellal Joseph. "Readmission with major abdominal complications after penetrating abdominal trauma." journal of surgical research 257 (2021): 69-78.

[2] Panwar, Yash, Sam Shan, Lily Owens, Chiu Kang, and Russell Hodgson. "Inefficient Admissions for Abdominal Pain Under an Acute General Surgical Unit." *World Journal of Surgery* 47, no. 10 (2023): 2401-2408.

[3] McCann, Aisling, Jan Sorensen, Deirdre Nally, Dara Kavanagh, and Deborah A. McNamara. "Discharge outcomes among elderly patients undergoing emergency abdominal surgery: registry study of discharge data from Irish public hospitals." BMC geriatrics 20, no. 1 (2020): 72.

[4] Colomier, Esther, Chloé Melchior, Joost P. Algera, Jóhann P. Hreinsson, Stine Störsrud, Hans Törnblom, Lukas Van Oudenhove et al. "Global prevalence and burden of meal-related abdominal pain." *BMC medicine* 20, no. 1 (2022): 71.

[5] Crowe, Sheila E. "Helicobacter pylori infection." *New England Journal of Medicine* 380, no. 12 (2019): 1158-1165.

[6] Go, M. F. "Natural history and epidemiology of Helicobacter pylori infection." *Alimentary pharmacology & therapeutics* 16 (2002): 3-15.

[7] Tamyalew, Yibeltal, Ayodeji Olalekan Salau, and Aleka Melese Ayalew. "Detection and classification of large bowel obstruction from X-ray images using machine learning algorithms." *International Journal of Imaging Systems and Technology* 33, no. 1 (2023): 158-174.

[8] Pham, Hieu H., Dung V. Do, and Ha Q. Nguyen. "Dicom imaging router: An open deep learning framework for classification of body parts from dicom x-ray scans." *arXiv preprint arXiv:2108.06490* (2021).

[9] Lakhani, Paras. "Deep convolutional neural networks for endotracheal tube position and X-ray image classification: challenges and opportunities." *Journal of digital imaging* 30 (2017): 460-468.

[10] Rehman, Arshia, and Fiaz Gul Khan. "A deep learning based review on abdominal images." *Multimedia Tools and Applications* 80, no. 20 (2021): 30321-30352.

[11] Kim, D. H., H. Wit, M. Thurston, M. Long, G. F. Maskell, M. J. Strugnell, D. Shetty, I.

M. Smith, and N. P. Hollings. "An artificial intelligence deep learning model for identification of small bowel obstruction on plain abdominal radiographs." *The British journal of radiology* 94, no. 1122 (2021): 20201407.

[12] Hussain, Emtiaz, Mahmudul Hasan, Md Anisur Rahman, Ickjai Lee, Tasmi Tamanna, and Mohammad Zavid Parvez. "CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images." *Chaos, Solitons & Fractals* 142 (2021): 110495.

[13] Weston, Alexander D., Panagiotis Korfiatis, Timothy L. Kline, Kenneth A. Philbrick, Petro Kostandy, Tomas Sakinis, Motokazu Sugimoto, Naoki Takahashi, and Bradley J. Erickson. "Automated abdominal segmentation of CT scans for body composition analysis using deep learning." *Radiology* 290, no. 3 (2019): 669-679.

[14] Cheng, Phillip M., and Harshawn S. Malhi. "Transfer learning with convolutional neural networks for classification of abdominal ultrasound images." *Journal of digital imaging* 30 (2017): 234-243.

[15] Bhattacharya, Samayan, Avigyan Bhattacharya, and Sk Shahnawaz. "Generating synthetic computed tomography (CT) images to improve the performance of machine learning model for pediatric abdominal anomaly detection." In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3865-3873. 2023.

[16] Ha, Tae Jun, Hee Sang Kim, Seong Uk Kang, DooHee Lee, Woo Jin Kim, Ki Won Moon, Hyun-Soo Choi et al. "Multi-classification of osteoporosis grading stages using abdominal computed tomography with clinical variables: application of deep learning with a convolutional neural network." *Journal of the Korean Society of Radiology* 18, no. 3 (2024): 187-201.

[17] Naz, Javeria, Muhammad Attique Khan, Majed Alhaisoni, Usman Tariq, and Seifedine Kadry. "Segmentation and Classification of Stomach Abnormalities Using Deep Learning." *Computers, Materials & Continua* 69, no. 1 (2021).

Study on Adjustable Stirrer Characteristics for Enhanced Field Uniformity in Electromagnetic Reverberation Chamber

Minhyuk Kim

Department of Electronic Engineering, Soonchunhyang University, Asan, 31538, Korea
Contact: MH.Kim@sch.ac.kr, phone +82-41-530-1326

Abstract—**This paper investigates the electromagnetic characteristics within an electromagnetic reverberation chamber by introducing an adjustable stirrer specifically designed to modify the shape of the reflectors. These modifications directly influence the internal electromagnetic boundary conditions, thereby impacting the field distribution inside the chamber. In conventional electromagnetic reverberation chambers, once stirrers are installed, the primary design parameters available for adjusting the field distribution are typically limited to tilting the antenna or repositioning it. However, by integrating an adjustable stirrer, it becomes possible to achieve greater control over the electromagnetic environment by dynamically altering the reflector shape. In this study, the performance of the adjustable stirrer was evaluated by comparing the uniformity of the electromagnetic field distribution with that achieved using a commonly employed Z-shaped stirrer. The results demonstrated that the adjustable stirrer significantly enhances field uniformity, providing a more flexible and effective method for optimizing electromagnetic conditions within the chamber.**

I. INTRODUCTION

To ensure seamless interoperability of electrical and electronic devices in electronic control systems, products must obtain electromagnetic compatibility (EMC) certification prior to being made available on the market. Presently, the automotive industry is undergoing a significant transition from internal combustion engines to environmentally-friendly automobiles, which include electric and hydrogen-powered ones. Additionally, a significant amount of research and development focuses on the implementation of new technologies, such as 5G, V2X, and autonomous driving. Consequently, electromagnetic compatibility (EMC) issues are becoming increasingly crucial in the automotive sector, which is shifting from mechanical to electronic products. The evaluation of EMC for automotive-related international standards (International Electrotechnical Commission (IEC) published in CISPR 16-1-4 [1], 16-2-3 [2], 16-2-4 [3] is performed in an electromagnetic semi-anechoic chamber (SAC).

However, for radiation immunity testing, a review and introduction of an electromagnetic reverberation chamber (RC) test method is underway to consider the deterioration of the propagation environment. In 2010, an international standard for the component level was published by the international organization for standardization (ISO) in ISO 11452-11 [4], followed by an international standard for the complete vehicle level in ISO 11451-5 [5]. Consequently, numerous domestic and foreign automobile manufacturers are either considering or implementing this test method.

Research into RCs began in the 1960s [6]. Unlike SACs, RCs don't require absorbers, making the equipment configuration for testing simple and eliminating the risk of fire. Additionally, it has the advantage of reducing evaluation time, although the calibration process for the test is relatively lengthy. Most significantly, unlike the SAC method of testing, which can only account for specific polarizations in certain directions, the immunity test can be carried out for arbitrary polarizations in all directions, leading to a more realistic assessment. The international standard for RC testing and measurement techniques published in IEC 61000-4-21 [7], states that the EMC evaluation of the equipment under test (EUT) in the working volume is required. This area guarantees uniformity of the electric field. Field uniformity (FU) is difficult to achieve at frequencies lower than the high frequency band. Typically, the design of RCs prioritizes high FU, a large working volume and the lowest usable frequency (LUF) for the same size. To meet these objectives, stirrers are used in RCs.

A stirrer is a tool that disrupts the electromagnetic field so that random polarization can take place in the RC, which is a cavity resonator having predefined modes. Although the Z-shape is popularly used, there are numerous stirrers such as propellers and intricate structures designed utilizing genetic algorithms [8-18]. Typically, the design variables that can alter the FU in the RC are extremely limited once a stirrer has been installed. These may include tilting or changing the position of the transmit antenna. If the FU of the working volume was not met during the calibration process, then the antenna parameters had to be adjusted and then recalibrated in all the frequency bands to be used. In IEC 61000-4-21 [7], it has been suggested that if the necessary uniformity is not achieved, the number of stirrers should be increased or the size of the stirrers should be enlarged. Depending on the chamber's size, this could pose a challenge. Additionally, various frequency bands necessitate distinct antennas, and different providers may have variable dimensions, which can impact the internal electromagnetic field. To address these constraints, there exist patents which permit uncomplicated modifications to the stirrer's structure post-installation. Previously, there have been studies conducted on adjusting the reflector angle, which necessitated the use of tools like wrenches. Due to the limited adjustability of the angle and poor

reproducibility, this technology faces restrictions on its applicability in practical settings.

In this paper, a 3D analysis tool was utilized for the objective verification of the patent. The tool analyzed the alteration of the electromagnetic field within the RC from the change in shape of the adjustable stirrer reflectors. Furthermore, an adjustable stirrer was constructed to investigate the effect of frequency on the change of FU. Then, the feasibility of using the adjustable stirrer was confirmed by benchmarking it against commercially available RC and stirrers, and comparing the results with the fabricated results.

II. FIELD UNIFORMITY

In IEC 61000-4-21, an experimental setup for electric field measurements to calculate FU in a RC is shown in Figure 1. The working volume should be at least $\lambda/4$ away from all metallic materials, including the walls of the RC, where λ is the wavelength at the LUF. FU is calculated using the measurements of the electric field components at each of the eight corners of the working volume.
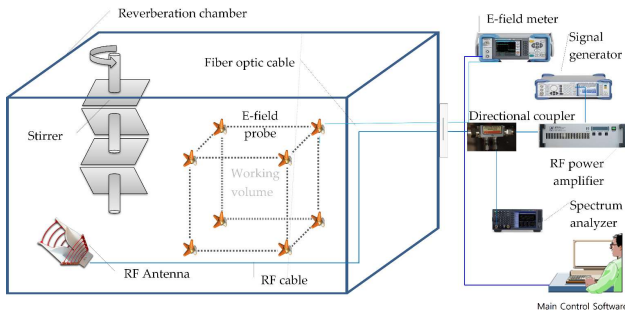

Fig. 1 A typical test setup for the field unifomity measurements as per IEC 61000-4-21.

III. RESULTS AND DISCUSSION

To verify the adjustment of the stirrer's FU, a similar-sized RC and stirrer were designed in line with commercial products. The analysis of the internal electromagnetic field was conducted with EMCoS studio software [19], a commercial software package. Figure 2 displays the simulation modelling.
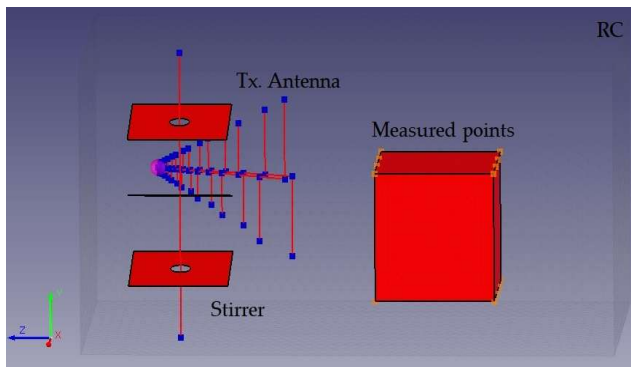

Fig. 2 Simulation modelling.

FU results were calculated by analyzing the electric field for 12 rotations of the adjustable stirrer in 30° increments

around the axis of rotation. The reflector was rotated in 15° increments around the central horizontal axis of rotation and the field was analyzed for three rotations. Figure 3 illustrates the standard deviation. The FU was determined by measuring the field at a total of 16 locations in order to verify the FU at more locations.
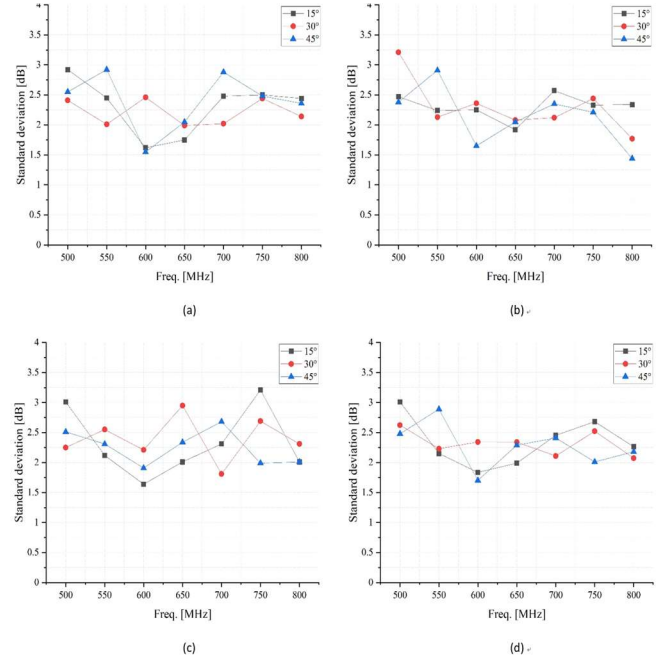

Fig. 3 Simulation results: (a) Ex, (b) Ey, (c) Ez, and (d) Et

The results of the analysis confirm the significant influence of the adjustable stirrer on the FU. To affirm the analysis, equivalent size RC and stirrer were produced and tested. The conventional Z-shaped stirrer (S_A) in Figure 4(a) was fabricated by benchmarking commercial products and used for result comparison with the adjustable stirrer.
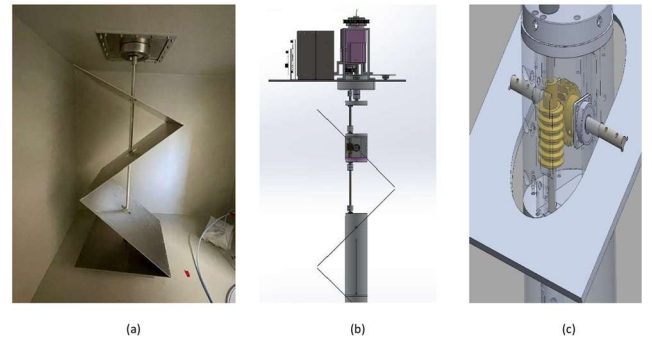

Fig. 4 Stirrers: (a) conventional z-shape, (b) design drawing, and (c) worm wheel of reflector operation part.

The adjustable stirrer (S_B) was constructed according to the design depicted in Figure 4(b), with each reflector being maneuvered by a worm wheel as shown in Figure 4(c). The reflectors are software-controlled, with movement in increments of 1 degree. To enable rotation, the reflector size is slightly smaller than that of S_A, while the same material as the S_A is used. Figure 5(a) displays a typical Z-shaped pattern, whereas Figure 5(b) shows an example of a stirrer with alterations made to the shape of the reflector.
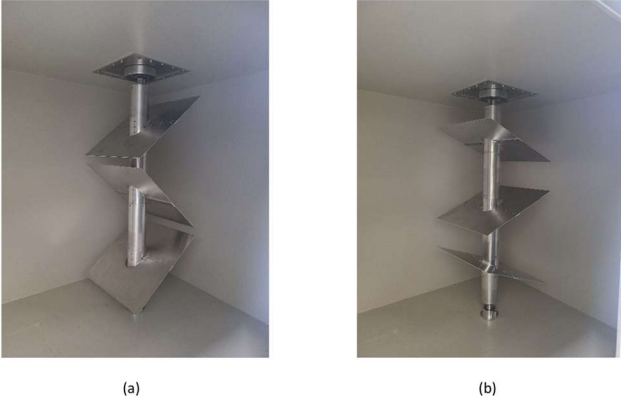
Fig. 5 Adjustable stirrer: (a) typical z-shape and (b) variation example.

To evaluate the changes in the internal electromagnetic field due to different stirrer shapes, the internal dimensions of the RC were kept in line with those of the simulation model and the commercial product. The commercial product's datasheet indicates that the LUF of the RC is 500 MHz at these dimensions. In this study, the field around the LUF was measured from 500 MHz to 800 MHz in 50 MHz steps.

In Figure 6, the red dashed line represents the standard's tolerance limit. The results displayed in Figure 6 indicate that neither stirrer meets the standard's accepted limits for at least one frequency. Especially, S_A exceeded the acceptable limits of the standard at 550 MHz in Figure 6(c). To comply with the standard's FU, S_A must undertake a new calibration at all frequencies by readjusting the antenna's position or the steering angle. On the other hand, in the trials involving S_B, there were no instances in which the outputs of every variant surpassed the allowable limits at any given frequency. Thus, modifying the angle of the stirrer reflector suffices to meet the standard requirements, and the time required for calibration decreases compared to that of S_A.
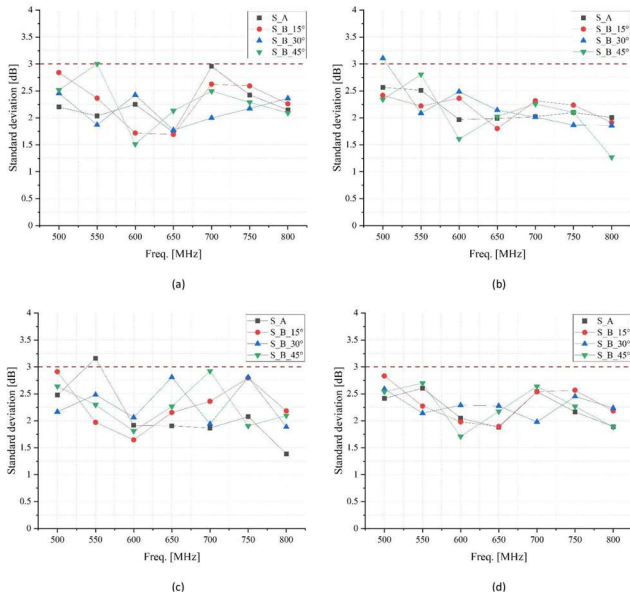

Fig. 6 Field uniformity comparison results for S_A and variations of S_B: (a) Ex, (b) Ey, (c) Ez, and (d) Et

Figure 7 illustrates the discrepancy between the simulations and experimental outcomes for each axis of the electric field.
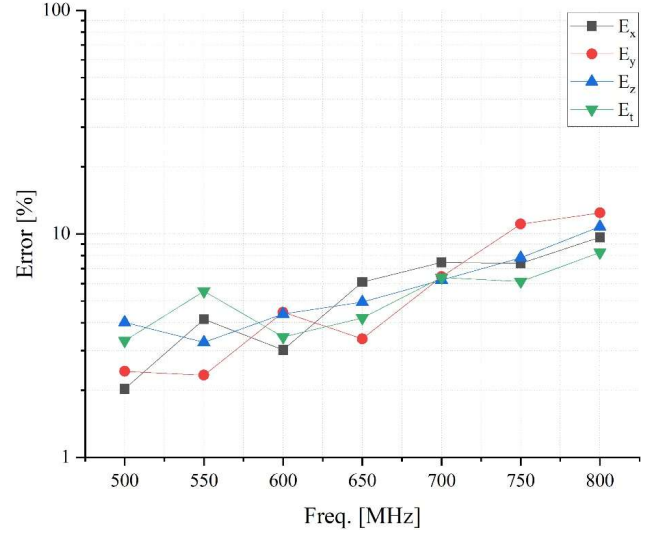

Fig. 7 Discrepancy between the simulations and experimental results.

IV. CONCLUSIONS

In this paper, we have assessed the efficacy of the patented adjustable stirrer for modifying the internal electromagnetic boundary conditions of an electromagnetic reverberation chamber (RC). Initially, we conducted a 3D numerical simulation using software to determine the impact of reflector shape alterations on field uniformity (FU). For our experiments, we constructed and evaluated a Z-shaped stirrer, which is widely used in RC applications. The FU was assessed by creating an adjustable stirrer of comparable size to the conventional stirrer. Testing indicated that the adjustable stirrer achieved the standard effortlessly by modifying the reflector angle, whereas the conventional stirrer failed to meet the acceptable limit at all frequencies and required recalibration. It is anticipated that the outcomes of this study will have wider practical implications beyond Z-shaped stirrers and extend to other variations of stirrers.

REFERENCES

[1] IEC (International Electrotechnical Commission), "Specification for Radio Disturbance and Immunity Measuring Apparatus and Methods – Part 1-4: Radio Disturbance and Immunity Measuring Apparatus – Antennas and Test Sites for Radiated Disturbance Measurements," *CISPR 16-1-4*, Ed. 4.0, 2019.

[2] IEC (International Electrotechnical Commission), "Specification for Radio Disturbance and Immunity Measuring Apparatus and Methods – Part 2-3: Methods of Measurement of Disturbances and Immunity – Radiated Disturbance Measurements," *CISPR 16-2-3*, Ed. 4.0, 2016.

[3] IEC (International Electrotechnical Commission), "Specification for Radio Disturbance and Immunity Measuring Apparatus and Methods Part 2-4: Methods of Measurement of Disturbances and Immunity – Immunity Measurements," *CISPR 16-2-4*, Ed. 1.0, 2003.

[4] ISO (International Organization for Standardization), "Road vehicles — Component test methods for electrical disturbances from narrowband radiated electromagnetic energy — Part 11: Reverberation chamber," *ISO 11452-11*, 2010.

[5]   ISO (International Organization for Standardization), "Road vehicles — Vehicle test methods for electrical disturbances from narrowband radiated electromagnetic energy — Part 5: Reverberation chamber," *ISO 11451-5*, 2023.

[6]   J. Yousaf, W. Nah, M. I. Hussein, J. G. Yang, A. ALTAF, and M. Elahi, "Characterization of Reverberation Chamber – A Comprehensive Review," *IEEE Access*, vol. 8, pp. 226591–226608, 2020. doi:10.1109/ACCESS.2020.3045028.

[7]   IEC (International Electrotechnical Commission), "Electromagnetic Compatibility (EMC) – Part 4-21: Testing and Measurement Techniques – Reverberation Chamber Test Methods," *IEC 61000-4-21*, Ed. 2.0, 2011.

[8]   K. Chen, Q. Xu, X. Shen, and C. Ren, "The Effect of Zigzag Boundaries on the Reverberation Chamber Performance," *IEEE Access*, vol. 9, pp. 145471–145476, 2021. doi:10.1109/ACCESS.2021.3123352.

[9]   J. Xu, Y. Zhao, and Q. Xu, "On the Time Domain Error Caused By the Frequency Domain Sampling Rate in a Reverberation Chamber," *IEEE Access*, vol. 7, pp. 78223–78227, 2019. doi:10.1109/ACCESS.2019.2921997.

[10]   T. Jia, Y. Huang, Q. Xu, Z. Tian, J. Jiang, and Q. Hua, "Frequency Domain Method for Scattering Damping Time Extraction of a Reverberation Chamber Based on Autocorrelation Functions," *IEEE Trans. Electromagn. Compat.*, vol. 62, pp. 2349–2357, 2020. doi:10.1109/TEMC.2020.2979014.

[11]   J. Tang, F. Li, J. Zheng, X. Chen, Y. Li, and J. Chen, "A New Mode Stirrer Design for the Reverberation Chamber," *Applied Computational Electromagnetics Society Journal*, vol. 36, pp. 1182–1188, 2021. doi:10.23907/ACES.2021.1182.

[12]   R. Serra, A. C. Marvin, F. Moglie, V. M. Primiani, A. Cozza, L. R. Arnaut, Y. Huang, M. O. Hatfield, M. Klingler, and F. Leferink, "Reverberation Chambers a la carte: An Overview of the Different Mode-Stirring Techniques," *IEEE Electromagnetic Compatibility Magazine*, vol. 6, pp. 63–78, 2017. doi:10.1109/MEMC.2017.7931986.

[13]   J. Clegg, A. C. Marvin, J. F. Dawson, and S. J. Porter, "Optimization of Stirrer Designs in a Reverberation Chamber," *IEEE Trans. Electromagn. Compat.*, vol. 47, pp. 824–832, 2005. doi:10.1109/TEMC.2005.86056.

[14]   C. Yu, Y. Zhao, H. Dai, W. Qi, F. Fang, W. Xia, and Q. Xu, "Improving the Reverberation Chamber Performance Using a Reconfigurable Source Stirring Antenna Array," *Electronics*, vol. 11, p. 1462, 2022. doi:10.3390/electronics11091462.

[15]   H. Sun, C. Gu, Z. Li, Q. Xu, J. Song, B. Xu, X. Dong, K. Wang, and F. Martín, "Enhancing the Number of Modes in Metasurfaced Reverberation Chambers for Field Uniformity Improvement," *Sensors*, vol. 18, p. 3301, 2018. doi:10.3390/s18103301.

[16]   H. Sun, C. Gu, Z. Li, Q. Xu, M. Wei, J. Song, B. Xu, X. Dong, K. Wang, and F. Martín, "Parametric Testing of Metasurface Stirrers for Metasurfaced Reverberation Chambers," *Sensors*, vol. 19, p. 976, 2019. doi:10.3390/s19040976.

[17]   A. Bakti, N. Kang, and J. Kwon, "Field Uniformity and Correlation Coefficient Analysis of KRISS Reverberation Chamber," *IEICE Trans. Commun.*, vol. E101-B, pp. 1745–, 2018. doi:10.1587/transcom.2017EBP3460.

[18]   J. Hong and C. Huh, "Optimization of stirrer with various parameters in reverberation chamber," *Progress In Electromagnetics Research*, vol. 104, pp. 15–?, 2010.

[19]   EMCoS Studio software, **EMCoS**, 2022.

A Hybrid Approach for Tumor Segmentation and Explain ability Using U-Net and Grad-CAM

Awais Khan[1], Yunyoung Nam[2] and Yongwon Cho[2]*

[1]Department of ICT Convergence, Soonchunhyang University, Asan 31538, Korea
[2]Department of Computer Science and Engineering, Soonchunhyang University, Asan, 31538, Korea
*Contact: dragon1won@sch.ac.kr

Abstract— **Bladder cancer is a prevalent and life-threatening disease where early detection and accurate bladder segmentation in computed tomography (CT) scans are critical for diagnosis, staging, and treatment planning. Traditional methods, such as deep convolutional neural networks with level sets (DCNN-LS), often require manual input and postprocessing, limiting their robustness in cases with poor image quality or tumor invasion. To address these challenges, we propose an automated U-Net-based deep learning model (U-DL) for bladder segmentation. Our dataset, comprising 296 cases collected at Anam Hospital, was split 80:20 for training and testing, with expert provided reference annotations. U-DL eliminates manual steps, improving efficiency and reliability. We evaluated model performance across different input resolutions and preprocessing strategies, including automated bladder cropping, using seven metrics: AVI, AMD, AVE, AHD, AAVE, Precision and Recall. Threshold analysis showed that the best performance occurred at 0.6, achieving a Dice Score of 0.95, IoU of 0.97, Precision of 0.98, and Recall of 0.97. The proposed framework improves diagnostic precision, reduces manual effort, and enhances bladder cancer detection.**

I. INTRODUCTION

Bladder cancer is one of the most frequently diagnosed cancers, contributing significantly to both morbidity and mortality rates. According to the American Cancer Society, approximately 81,190 new cases were projected in 2018 in the United States, with 62,380 cases among men and 18,810 among women, alongside an estimated 17,240 deaths, including 12,520 men and 4,720 women [1].

Multidetector row computed tomography (MDCT) urography is widely recognized as the most effective imaging modality for evaluating the urinary tract, utilizing a combination of unenhanced, corticomedullary, nephrographic, and excretory phase imaging [2,3]. MDCT urography (CTU) can assess multiple structures, including the kidneys, the intrarenal collecting systems, and the ureters [4,5]. Each CTU examination typically generates at least 300 reconstructed slices with slice intervals of approximately 1.25 mm, ranging from 200 to 600 slices per case. These scans may reveal multiple lesions and a variety of urinary tract abnormalities, requiring considerable effort and time for radiologists to accurately interpret the findings. Consequently, disease detection rates can vary significantly;

for instance, reported sensitivity for detecting bladder cancer ranges between 59% and 92% [6,7].

Several studies have focused on bladder segmentation across different imaging modalities. In the case of magnetic resonance imaging (MRI), multiple level set-based approaches have been proposed to segment the bladder walls. Duan et al. [8] employed a coupled level set method to segment the bladder wall on T1-weighted MR images in six patients. Chi et al. [9] utilized a geodesic active contour model to segment the inner bladder wall on T2-weighted images and applied a maximum wall thickness constraint on T1-weighted images to segment the outer wall in a cohort of 11 patients. Han et al. [10] developed an adaptive Markov random field model combined with coupled level set information to perform bladder wall segmentation on T1-weighted MR images in six patients.

Similarly, Qin et al. [11] introduced an adaptive shape prior-constrained level set technique for segmenting the bladder wall on T2-weighted images in 11 patients. Although effective, these level set methods are typically time-intensive and face challenges in defining an appropriate stopping criterion. To address efficiency concerns, Xu et al. [12] proposed a continuous max-flow framework utilizing global convex optimization for bladder segmentation on T2-weighted images in five patients. Despite their advancements, these methods were specifically designed for MR images and were validated only on relatively small datasets.

Regarding computed tomography (CT) imaging, Chai et al. [13] proposed a semi-automatic bladder segmentation method based on statistical shape modeling, tested in 23 patients. Additionally, Hadjiiski et al. [14] developed the Conjoint Level Set Analysis and Segmentation System (CLASS) for bladder segmentation in 81 CTU examinations. However, both approaches showed limitations, including a strong dependency on initial conditions and validation on limited sample sizes.

Deep convolutional neural networks (DCNNs) have recently emerged as a powerful tool, achieving notable success in classifying natural images when trained on large datasets [15]. The application of convolutional neural networks (CNNs) for recognizing patterns in medical images dates back to the 1990s [16].

A. Major Challenges

Nevertheless, achieving precise bladder segmentation in CTU images remains challenging. In the excretory phase, the bladder can present regions with and without excreted contrast material, complicating the differentiation between the bladder wall and surrounding soft tissues, especially when contrast opacification is incomplete. Moreover, variability in bladder shapes, sizes, and the presence of abnormalities further increases the difficulty of obtaining accurate segmentation results.

B. Major Contributions

The objective of this study is to develop a fully automated, robust deep learning framework for bladder segmentation in CT scans without the need for manual input or postprocessing. The proposed framework includes the following steps:

- The dataset of 296 cases, annotated by expert radiologists, was split into 80% training and 20% testing sets.
- Input CT scans were preprocessed using automated bladder cropping and different input resolutions to enhance model performance.
- A U-Net-based deep learning model (U-DL) was trained for bladder segmentation, eliminating manual user input and additional postprocessing.
- The segmentation performance was quantitatively evaluated using eight metrics: AVI, AMD, AVE, AHD, AAVE, Precision, Recall, and AJI.
- Threshold analysis was performed to optimize segmentation accuracy.
- Grad-CAM was applied to visualize the U-Net model's decision-making by generating heatmaps of influential regions in the input images.

II. METHODOLOGY

In this study, we propose a novel deep learning-based approach for automated bladder segmentation in CT scans, as illustrated in Figure 1. The methodology consists of several key steps: data preprocessing through automated bladder cropping and input resizing, segmentation using a U-Net-based model (U-DL), and evaluation using seven quantitative metrics. The U-DL model was trained on 296 CT cases annotated by expert radiologists. The segmentation performance was optimized by analyzing different thresholds, with the final model achieving high precision and robustness without requiring manual input or postprocessing.
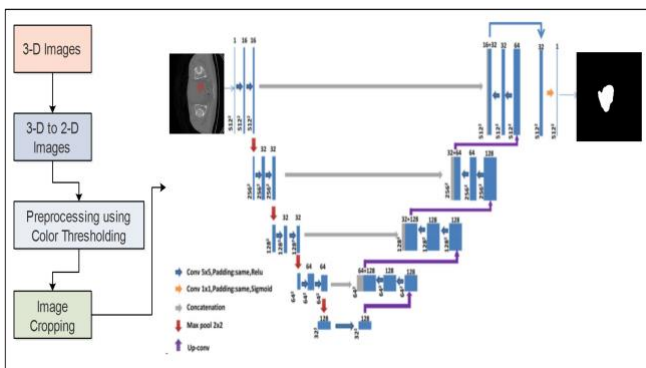

Figure 1. Proposed architecture of bladder cancer detection.

A. Data Collection

The dataset used in this study consists of CT scan images collected from Korea University Anam Hospital. Initially, the dataset comprised three-dimensional (3D) CT scan volumes, which were converted into two-dimensional (2D) image slices using ITK-SNAP software to facilitate easier processing. A total of 296 2D CT images were extracted, each containing regions that highlight the tumor areas. Along with these images, 296 corresponding mask images were generated to accurately delineate the tumor boundaries. Each mask precisely marks the tumor region within the respective CT slice. The dataset was organized to maintain a one-to-one correspondence between the images and their masks. These labeled images were later used for training and testing purposes. A sample from the dataset is shown below.
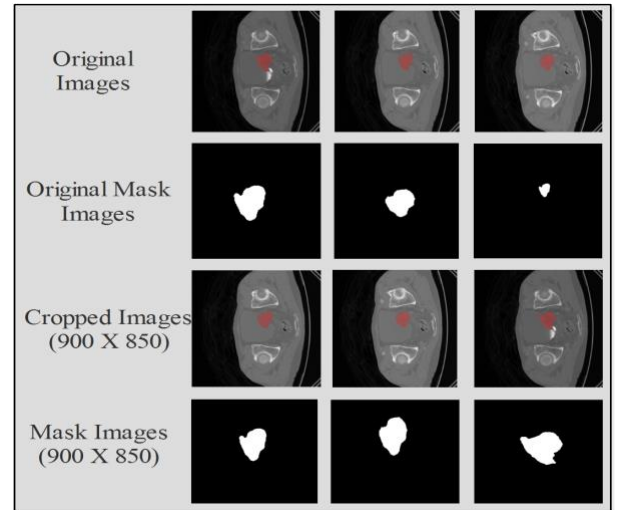

Figure 2. Sample of proposed bladder dataset.

B. Bladder segmentation using U-Net model

In this study, the segmentation techniques are built upon the U-Net neural network architecture. The models were implemented using Keras with TensorFlow as the backend. To optimize the segmentation performance for bladder images, we modified and fine-tuned the original U-Net structure and its parameters. For simplicity, we refer to this customized version as U-Net Deep Learning (U-DL) throughout our study. U-DL models were developed and evaluated to segment the bladder effectively. Specifically, we designed and compared: (a) 2D U-DL models, using 2D CT slices volumes as input, respectively; (b) U-DL models trained with CT images at varying resolutions; and (c) U-DL models with and without an automated bladder cropping step applied during image preprocessing. Proposed architecture of U-Net model is shown in Figure. 3.
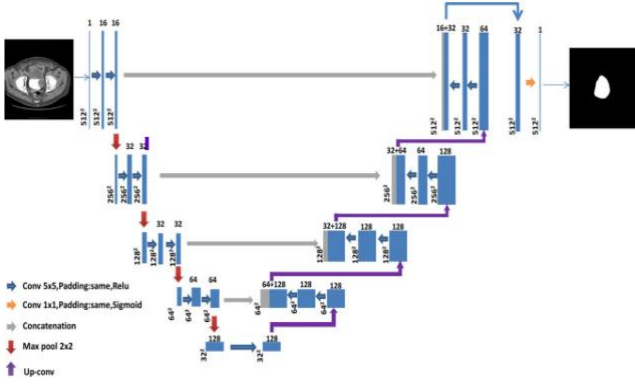
Figure 3. Proposed architecture of U-Net deep learning model.

C. 2-Dimensional U-Net model

For each CTU case, a 2D U-DL model was configured based on a 2D U-Net architecture to predict a bladder likelihood map slice by slice. The 2D U-DL was trained using the 2D slices extracted from the training dataset. To reduce processing time, slices without the bladder were excluded by selecting the approximate first and last slices containing the bladder, as identified by a radiologist. All remaining slices with bladder content were directly used as inputs to the 2D U-DL model, eliminating the need for a bounding box around the bladder. The model was trained using corresponding manually segmented bladder masks as target outputs for each input slice. Examples of input slices and their associated target masks are shown in Figure 2.

The architecture of the 2D U-DL is depicted in Figure 3. It features a contracting path that captures contextual information using multiresolution features and an expanding path that reconstructs the bladder region. In the contracting path, two 5×5 padded convolutions are applied repeatedly, each followed by a ReLU activation and a 2×2 max pooling operation with a stride of 2 for downsampling. Same padding was employed to maintain the spatial dimensions across layers, which differentiates it structurally from the traditional U-Net. In the expanding path, feature maps are upsampled, followed by a 5×5 padded convolution and ReLU activation (up-convolution). Additionally, two more 5×5 padded convolutions with ReLU activation are applied, and feature maps from the contracting path are concatenated to improve localization accuracy. This design choice, where concatenation happens after two convolutional layers rather than before, also sets it apart from standard U-Net. Finally, a 1×1 convolution layer with sigmoid activation maps the multichannel feature vector into a per-pixel probability representing the likelihood of belonging to the bladder region.

D. Bladder segmentation using U-Net model

To assess the impact of image resolution on U-DL-based bladder segmentation, we trained and compared U-DL models using input images of different resolutions. The original CT slices were downsampled to 900×850 pixels by applying a 2×2 box filter followed by downsampling with a stride of 2, effectively doubling the pixel size.

Additionally, to investigate the influence of surrounding background on segmentation performance, an automated cropping preprocessing step was designed. We compared the U-DL models trained with and without this cropping procedure. In the preprocessing step, a cantered square

region was automatically extracted from each slice. For the original images, a 900×850 pixel region was cropped, while for the downsampled 900×850 images, pixel region was selected. The cropped regions included the entire bladder while minimizing surrounding background.

An illustration of the downsampling and cropping process is presented in Figure 4. As a result of these operations, images with varying fields of view and sizes were generated for training and evaluation.
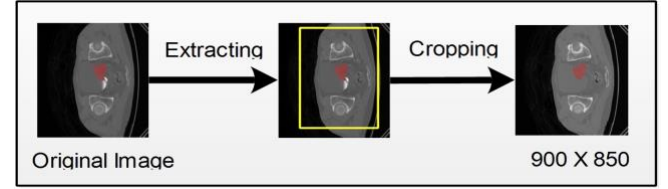

Figure 4. An example of the downsampling and cropping process.

III. RESULTS AND DISCUSSION

In this section, we present the results obtained using the proposed U-DL framework for bladder segmentation. The dataset, comprising 296 cases, was divided into an 80:20 ratio, with 80% used for training and 20% for testing. To compare model performance, experiments were conducted using two input resolutions: the original CT scan size and resized images of 900×850 pixels, generated through automated bladder cropping, as illustrated in Figure. 2. The output likelihood maps were resized to match the corresponding input dimensions. The training process involved selecting key hyperparameters, including a learning rate of 0.0001, 100 epochs, and a mini-batch size of 4 for the 2D U-DL models. Training was performed using the Adam optimizer with a binary cross-entropy loss function. Minor adjustments to the learning rate did not significantly impact model performance, and convergence was consistently achieved across different configurations.

All experiments were conducted using Python on a desktop computer equipped with 64 GB RAM, a 2 TB SSD, and a 16 GB NVIDIA GeForce RTX 4070 Ti Super GPU. The segmentation performance was quantitatively assessed using comprehensive evaluation metrics, including AVI, AMD, AVE, AHD, AAVE, Precision, Recall, and AJI, providing insights into the accuracy and robustness of the proposed model.

C. Proposed results of U-Net model on original input image dataset

In this subsection, the segmentation results for bladder CT scans were evaluated by analyzing the effect of varying decision thresholds. Multiple metrics were computed across thresholds ranging from 0.1 to 0.9, including Dice Score, Intersection over Union (IoU), Absolute Volume Intersection (AVI), Absolute Volume Error (AVE), Recall, and Precision. The detailed segmentation performance across thresholds is presented in Table 1. The results show that increasing the decision threshold generally improved the Dice Score and IoU. Specifically, the Dice Score increased from 0.88 at a threshold of 0.1 to a maximum of 0.96 at a threshold of 0.9. Similarly, IoU improved from 0.93 to 0.98 over the same range. AVI and AVE values indicated improved segmentation accuracy up to thresholds

of 0.5–0.6, after which a slight degradation in performance was observed.

The optimal performance, balancing all metrics, was achieved at a threshold of 0.5. At this point, the Dice Score was 0.92, IoU was 0.96, AVI reached 94.93%, AVE was minimized to 0.50%, Recall was 0.97, and Precision was 0.97. Additionally, while Recall slightly decreased at higher thresholds, Precision increased, reaching 0.99 at a threshold of 0.9. This trend indicates a trade-off between sensitivity and specificity, where higher thresholds favor greater precision at the cost of reduced recall.

Table 1. Results of original images.

Threshold	Dice Score	IoU	AVI (%)	AVE (%)	Recall	Precision
0.1	0.88	0.93	91.26	8.53	0.99	0.91
0.2	0.88	0.94	93.21	5.56	0.99	0.93
0.3	0.90	0.95	94.23	3.62	0.98	0.95
0.4	0.91	0.95	94.69	1.96	0.98	0.96
0.5	0.92	0.96	94.93	0.50	0.97	0.97
0.6	0.93	0.96	94.74	0.90	0.96	0.97
0.7	0.93	0.97	93.30	2.61	0.95	0.98
0.8	0.94	0.97	93.04	4.79	0.94	0.98
0.9	0.96	0.98	90.76	8.25	0.91	0.99

D. Proposed results of U-Net model on 900 X 850 cropped images

To determine the optimal decision threshold for binary segmentation, model performance was evaluated across thresholds ranging from 0.1 to 0.9, as summarized in Table 2.

The results demonstrated a stable improvement in segmentation performance with increasing thresholds. Specifically, the Dice Score improved from 0.90 at a threshold of 0.1 to a maximum of 0.98 at 0.9, while IoU increased from 0.94 to 0.99. AVI also improved, reaching its peak value of 95.85% at a threshold of 0.6. Concurrently, AVE decreased significantly with higher thresholds, achieving a minimum value of 0.28% at a threshold of 0.6, indicating reduced volumetric error. Recall remained consistently high (>0.95) across all thresholds but gradually decreased from 0.99 to 0.92 as the threshold increased. In contrast, Precision improved steadily, reaching 0.99 at thresholds 0.8 and 0.9. This behavior reflects the typical trade-off between Recall and Precision, where increasing the threshold enhances specificity at the cost of a slight reduction in sensitivity.

The best balance between high Dice Score, IoU, Recall, and Precision was achieved between thresholds of 0.5 and 0.6. Notably, at a threshold of 0.6, the Dice Score was 0.95, IoU was 0.97, AVI was 95.85%, AVE was minimized to 0.28%, Recall was 0.97, and Precision was 0.98. Figure 5

shows an example of the bladder segmentation results using U-DL from a case in the test set.

Table 2. Results of cropped images (900 X 850).

Threshold	Dice Score	IoU	AVI (%)	AVE (%)	Recall	Precision
0.1	0.90	0.94	92.12	8.14	0.99	0.92
0.2	0.92	0.95	94.14	5.46	0.99	0.94
0.3	0.92	0.96	95.09	3.77	0.99	0.95
0.4	0.93	0.96	95.61	2.44	0.98	0.96
0.5	0.94	0.97	95.80	1.12	0.98	0.97
0.6	0.95	0.97	95.85	0.28	0.97	0.98
0.7	0.95	0.98	95.56	1.92	0.96	0.98
0.8	0.96	0.98	94.50	3.93	0.95	0.99
0.9	0.98	0.99	92.41	6.69	0.92	0.99

E. Visualization of model prediction using Grad-Cam

In this study, we use Grad-CAM (Gradient-weighted Class Activation Mapping) to visualize the decision-making process of the U-Net model for image segmentation. Grad-CAM generates heatmaps that highlight the regions in the input image most influential for the model's prediction. We preprocess the image by removing red regions, resizing it, and normalizing it before feeding it into the model. The Grad-CAM algorithm computes gradients of the output with respect to a specific convolutional layer and creates a heatmap that emphasizes the areas with the highest contribution to the prediction. This heatmap is then overlaid on the original image, providing valuable insights into the model's focus areas and enhancing the interpretability of its predictions. By incorporating Grad-CAM, we aim to offer a deeper understanding of the model's performance beyond traditional evaluation metrics. Visualization of bladder detection is shown in Figure. 5.
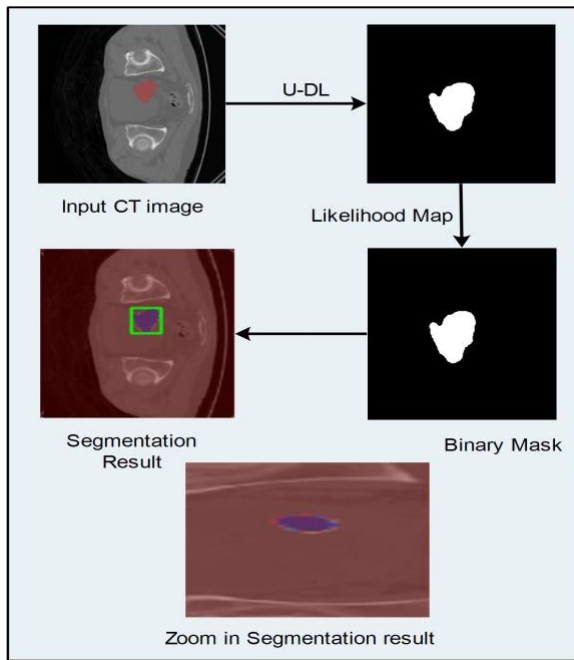
Figure 5. Visualization of the bladder segmentation results using two-dimensional (2D) U-DL with 900×850 pixel resolution as input and with automated cropping in preprocessing from a case in the test set. The lighter (blue) contour represents segmentation using 2D U-DL with threshold 0.6.

IV. CONCLUSIONS

In future work, we aim to enhance the quality and accuracy of bladder cancer detection by applying various image enhancement techniques. Additionally, we will implement and evaluate multiple models to compare their performance, with the goal of improving both segmentation and detection accuracy. To further strengthen the robustness of the model, we plan to expand the dataset by increasing the number of images, thereby allowing for more comprehensive training and validation. Furthermore, we will incorporate explainable AI techniques such as Grad-CAM++ and Score-CAM to provide better interpretability and insights into the model's decision-making process.

ACKNOWLEDGMENT

REFERENCES

[1]   American Cancer Society. Key Statistics for Bladder Cancer; 2018. www.cancer.org

[2]   Caoili EM, Cohan RH, Korobkin M, et al. Urinary tract abnormalities: initial experience with multi–detector row CT urography. Radiology. 2002;222:353–360.

[3]   Gupta R, Raghuvanshi S. Multi–detector CT urography in the diagnosis of urinary tract abnormalities. Int J Med Res Rev. 2016;4:222–226.

[4]   Akbar SA, Mortele KJ, Baeyens K, Kekelidze M, Silverman SG. Multidetector CT urography: techniques, clinical applications, and pitfalls. Semin Ultrasound CT MRI. 2004;25:41–54.

[5]   Noroozian M, Cohan RH, Caoili EM, Cowan NC, Ellis JH. Multislice CT urography: state of the art. Br J Radiol. 2004;77:S74–S86.

[6]   Park SB, Kim JK, Lee HJ, Choi HJ, Cho K-S. Hematuria: portal venous phase multi–detector row CT of the bladder—a prospective study. Radiology. 2007;245:798–805.

[7]   Sudakoff GS, Dunn DP, Guralnick ML, Hellman RS, Eastwood D, See WA. Multidetector computerized tomography urography as the primary imaging modality for detecting urinary tract neoplasms in patients with asymptomatic hematuria. J Urol. 2008;179:862–867.

[8]   Duan C, Liang Z, Bao S, et al. A coupled level set framework for bladder wall segmentation with application to MR cystography. IEEE Trans Med Imaging. 2010;29:903–915.

[9]   Chi JW, Brady M, Moore NR, Schnabel JA. Segmentation of the bladder wall using coupled level set methods. Paper presented at: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro; 30 March-2 April 2011, 2011.

[10] Han H, Li L, Duan C, Zhang H, Zhao Y, Liang Z. A unified EM approach to bladder wall segmentation with coupled level-set constraints. Med Image Anal. 2013;17:1192–1205.

[11] Qin X, Li X, Liu Y, Lu H, Yan P. Adaptive shape prior constrained level sets for bladder MR image segmentation. IEEE J Biomed Health Inform. 2014;18:1707–1716.

[12] Xu X-P, Zhang X, Liu Y, et al. Simultaneous segmentation of multiple regions in 3D bladder MRI by efficient convex optimization of coupled surfaces. In: Zhao Y, Kong X, Taubman D, eds. Image and Graphics. Cham: Springer; 2017.

[13] Chai X, Herk MV, Betgen A, Hulshof M, Bel A. Automatic bladder segmentation on CBCT for multiple plan ART of bladder cancer using a patient-specific bladder model. Phys Med Biol. 2012;57:3945.

[14] Hadjiiski L, Chan H-P, Cohan RH, et al. Urinary bladder segmentation in CT urography (CTU) using CLASS. Med Phys. 2013;40:111906.

[15] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Comm ACM. 2012;60:84–90.

[16] Lo SB, Lou SA, Jyh-Shyan L, Freedman MT, Chien MV, Mun SK. Artificial convolution neural network techniques and applications for lung nodule detection. IEEE Trans Med Imaging. 1995;14:711–718.