Ihsan Ullah,PhD

Department of Robotics and Mechatronics Engineering,
Daegu Gyeongbuk Institute of Science and Technology (DGIST),
Daegu, South Korea.
Ihsanullah736@gmail.com / Ihsankhan@dgist.ac.kr
https://ihsan149.github.io
https://sites.google.com/view/mispl/home
https://sites.google.com/view/ccg-dgist/home
(+82) 10-2291-5091

Research Statement

Most deep learning-based methods require a considerable amount of labeled data, which is difficult to come by in the computer vision and medical field. I am interested in developing unsupervised domain adaptation methods to learn from synthetic data to mitigate labeling efforts. Furthermore, my research interests also include combining robotics and AI to develop intelligent robotic systems capable of automating processes and making autonomous decisions (see page 7).
In addition to my research interests, I am also deeply invested in teaching and mentoring. I believe that the classroom is an excellent platform to introduce and explore the complex topics of Computer Science/Robotics and AI (see page 12).

Employment History

- 2023-04 – In Progress: **Post-Doctoral Researcher,** CCG, Division of Intelligent Robot, DGIST.
- 2018-03 – 2023-02: **PhD Research Student,** MISPL, Robotics and Mechatronics Engineering, DGIST.
- 2017-09 – 2018-02: **Post MS Researcher,** MISPL, Robotics and Mechatronics Engineering, DGIST.
- 2015 – 2017-08: **MS Research Student,** Software Engineering Lab, Jeonbuk National University .

Education

- 2018 – 2023: **Ph.D. Robotics and Mechatronics Engineering,** Daegu Gyeongbuk Institute of Science and Technology.
  Thesis title: *Catheter Localization and Tracking using Convolutional Neural Networks with Generative Modeling.*
- 2015 – 2017: **M.S. Computer Science and Engineering,** Jeonbuk National University.
  Thesis title: *Vehicles Information Extraction Based on Deep Neural Network.*
- 2010 – 2014: **B.S. Computer Science,** Abdul Wali Khan University Mardan.
  Thesis title: *Developed Dynamic Website (Crime File Management System).*

Research Publications

Representative Articles

1. **Ullah, I.**, An, S., Kang, M., Chikontwe, P., Lee, H., Choi, J., & Park, S. H. (2024a). *Video domain adaptation for semantic segmentation using perceptual consistency matching.* doi:https://doi.org/10.1016/j.neunet.2024.106505
2. **Ullah, I.**, Chikontwe, P., Choi, H., Yoon, C.-H., & Park, S. H. (2021a). *Synthesize and segment: Towards improved catheter segmentation via adversarial augmentation.* doi:10.3390/app11041638

3. **Ullah, I.**, Chikontwe, P., & Park, S. H. (2019a). *Catheter synthesis in x-ray fluoroscopy with generative adversarial networks*. Springer. Retrieved from `https://link.springer.com/chapter/10.1007/978-3-030-32281-6_13`

Articles in Submission/Preparation

1. Ali, H., **Ullah, I.**, Asim, W., Adeeb, S., & Lee, H. (In Preparation). White matter multiple sclerosis lesion segmentation under distributional shifts. In *Multiple sclerosis journal*. Sage.
2. **Ullah, I.**, & Lee, H. (In Preparation). Towards real-time drone detection and tracking in thermal infrared videos for anti-drone system. In *Applied intelligence*. Springer.

Journal Articles

1. Ho-Gun, H., Deokgi, J., **Ullah, I.**, Junichi, T., Jaesung, H., & Hyunki, L. (2024). Target-specified reference-based deep learning network for joint image deblurring and resolution enhancement in surgical zoom lens camera calibration. *Computer Methods and Programs in Biomedicine*.
2. **Ullah, I.**, An, S., Kang, M., Chikontwe, P., Lee, H., Choi, J., & Park, S. H. (2024b). Video domain adaptation for semantic segmentation using perceptual consistency matching. *Neural Networks*, 106505. doi:`https://doi.org/10.1016/j.neunet.2024.106505`
3. **Ullah, I.**, Baber, S., Ali, F., El-Sappagh, S., Abuhmed, T., & Park, S. H. (2023). A deep learning based dual encoder-decoder framework for anatomical structure segmentation in chest fluoroscopic images. *Scientific Reports*, *13*(1), 791. doi:`https://doi.org/10.1038/s41598-023-27815-w`
4. Jeong, J., Hong, S. T., **Ullah, I.**, Kim, E. S., & Park, S. H. (2022). Classification of the confocal microscopy images of colorectal tumor and inflammatory colitis mucosa tissue using deep learning. *Diagnostics*, *12*(2), 288. doi:`10.3390/diagnostics12020288`
5. Shoaib, M., Hussain, T., Shah, B., **Ullah, I.**, Shah, S. M., Ali, F., & Park, S. H. (2022). Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease. *Frontiers in plant science*, *13*, 1031748. doi:`10.3389/fpls.2022.1031748`
6. Awan, N., Khan, S., Rahmani, M. K. I., Tahir, M., Alam, N., Alturki, R., & **Ullah, I.** (2021). Machine learning-enabled power scheduling in iot-based smart cities. *Comput. Mater. Contin*, *67*(2), 2449–2462.
7. **Ullah, I.**, Chikontwe, P., Choi, H., Yoon, C.-H., & Park, S. H. (2021b). Synthesize and segment: Towards improved catheter segmentation via adversarial augmentation. *Applied Sciences*, *11*(4), 1638. doi:`10.3390/app11041638`
8. Allan, M., Kondo, S., **Ullah, I.**, Bodenstedt, S., Leger, S., Kadkhodamohammadi, R., … Pedersen, M. et al. (2020). 2018 robotic scene segmentation challenge. *arXiv preprint arXiv:2001.11190*. Retrieved from `https://arxiv.org/abs/2001.11190v3`
9. Lee, H. J., **Ullah, I.**, Wan, W., Gao, Y., & Fang, Z. (2019). Real-time vehicle make and model recognition with the residual squeezenet architecture. *Sensors*, *19*(5), 982. doi:`10.3390/s19050982`
10. **Ullah, I.**, Chikontwe, P., & Park, S. H. (2019d). Real-time tracking of guidewire robot tips using deep convolutional neural networks on successive localized frames. *IEEE Access*, *7*, 159743–159753. doi:`10.1109/ACCESS.2019.2950263`

Conference Proceedings

1. Chikontwe, P., **Ullah, I.**, Kim, J., Won, D., & Park, S. H. (2019). Recurrent attention models for tissue histopathology image classification. Retrieved from `http://ipiu.or.kr/2019/oral.html`
2. **Ullah, I.**, Chikontwe, P., & Park, S. H. (2019b). Catheter synthesis in x-ray fluoroscopy with generative adversarial networks. In *"MICCAI PRIME", workshop* (pp. 125–133). Springer. Retrieved from `https://link.springer.com/chapter/10.1007/978-3-030-32281-6_13`

3 **Ullah, I.**, Chikontwe, P., & Park, S. H. (2019c). Guidewire tip tracking using u-net with shape and motion constraints. In *2019 international conference on artificial intelligence in information and communication (ICAIIC)* (pp. 215–217). doi:10.1109/ICAIIC.2019.8669088

4 Alam, M. N., **Ullah, I.**, & Al-Absi, A. A. (2018). Deep learning-based apple defect detection with residual squeezenet. In *International conference on smart computing and cyber security: Strategic foresight, security challenges and innovation* (pp. 127–134). Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-981-15-7990-5_12

5 **Ullah, I.**, & Lee, H. J. (2017). Moving vehicle detection and information extraction based on deep neural network. In *Proceedings of the international conference on image processing, computer vision, and pattern recognition (IPCV)* (pp. 102–107). The Steering Committee of The World Congress in Computer Science, Computer … Retrieved from https://csce.ucmss.com/books/LFS/CSREA2017/IPC6015.pdf

6 **Ullah, I.**, & Lee, H. J. (2016a). An approach of locating korean vehicle license plate based on mathematical morphology and geometrical features. In *2016 international conference on computational science and computational intelligence (csci)* (pp. 836–840). IEEE. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7881455

7 **Ullah, I.**, & Lee, H. J. (2016b). An effective algorithm for shadow removal from moving vehicles. In *International symposium on information technology convergence (ISITC 2016), shanghai china*. Retrieved from https://www.researchgate.net/publication/317184010_An_Effective_Algorithm_for_Shadow_Removal_from_Moving_Vehicles_Based_on_Morphology

8 **Ullah, I.**, & Lee, H. J. (2016c). License plate detection based on rectangular features and multilevel thresholding. In *Proceedings of the international conference on image processing, computer vision, and pattern recognition (IPCV)* (p. 153). The Steering Committee of The World Congress in Computer Science, Computer … Retrieved from http://worldcomp-proceedings.com/proc/p2016/IPC4045.pdf

9 **Ullah, I.**, & Lee, H. J. (2016d). Moving object detection based on background subtraction. In *Conference of KIISE, june 29-july 01, 2016, jeju, south korea*. Retrieved from https://www.dbpia.co.kr/pdf/pdfView.do?nodeId=NODE07017673

Patents

1. Tracking of Catheter Tip using Convolutional Neural Network

Title in Korean:	합성곱신경망을 이용한 카테터 끝단 추적기법
Inventors:	**Ihsan Ullah**, DongKyu Won, Sang Hyun Park
Patent Number:	10-2174246-0000
Application Number:	10-2018-0054790
Registration Date:	29/10/2020

2. Catheter Synthesis in X-Rays

Title in Korean:	X-ray 카테터 시술 영상합성 기법
Inventors:	**Ihsan Ullah**, Philip Chikontwe, Sang Hyun Park
Patent Number:	10-2530103-0000
Application Number:	10-2020-0006741
Registration Date:	27/ 04/ 2023

3. Method and Apparatus for Learning Domain Adaptation in Video Segmentation (Under-review)

Title in Korean:	영상 분할에서의 도메인 적응 학습 방법 및 장치
Inventors:	**Ihsan Ullah**, Sang Hyun Park
Filing Date:	2023-01-18

4. AI-based Drone Tracking System (Under-review)

Title in Korean: 인공지능 기반 드론 추적 시스템

Inventors: **Ihsan Ullah**, Hyunki Lee, Roh Jung-Gu

Filing Date: 2024-07-18

5. Drowning Detection and Monitoring based on Camera Images and Rescue Systems using Unmanned Drones (Under-review)

Title in Korean: 카메라 영상 기반 익사자 감지와 모니터링 및 무인 드론을 활용한 익사자 구출 시스템

Inventors: **Ihsan Ullah**, Hyunki Lee, Gyumin Park

Filing Date: 2024-08-28

Participation in MICCAI Challenges

2018
- **MICCAI- EndoVis:** The challenge focuses on robotic surgical scene segmentation.
  Title: *2018 Robotic Scene Segmentation Challenge.*

2024
- **MICCAI- STIR Challenge 2024:** The STIR Challenge was designed to quantify methods in deformable tracking, mapping and reconstruction.
  Title: *Track points accurately and efficiently in videos.* ***Runner-up***
- **MICCAI- SegSTRONG-C:** To evaluate algorithm robustness against non-adversarial corruptions.
  Title: *SegSTRONG-C: Segmenting Surgical Tools Robustly On Non-adversarial Generated Corruptions – An EndoVis'24 Challenge.* ***6th/40 registered teams***

Teaching & Supervision

Visiting-Lecturer

Grad & UG Course
- Visiting-Lecturer for **Applied AI for Robotic Systems (IE803, 2024 Spring)**.

Course Description
- This course provided an in-depth exploration of Applied Artificial Intelligence (AI) in intelligent robotic systems. Through lectures, exercises, and projects, students gained an understanding of AI integration in robotics. The curriculum covered machine learning, computer vision, and AI techniques for robotics, focusing on applications such as recognition, segmentation, classification, detection, and tracking.

Responsibilities
- Designed this course for graduate & under-grad students.
- Grading assignments/exams and providing feedback to students.
- Delivering lectures.
- Conducting lab experiments to help student.

Grad & UG Course
- Visiting-Lecturer for **Applied AI for Robotic Systems (IE803, 2025 Spring)**.

Course Description
- This course explores robotics fundamentals and applications, covering kinematics, dynamics, control, and ROS. Students engage in hands-on activities like kinematics problem-solving, PID tuning, and trajectory generation. Advanced topics include motion planning with UR10 and integrating machine learning (e.g., CNNs) for classification, detection, and tracking in robotics.

Responsibilities
- Designed this course for graduate & under-grad students.
- Grading assignments/exams and providing feedback to students.
- Delivering lectures.
- Conducting lab experiments to help students.

Teaching Assistance(TA)

Gradute Course
- TA for **Deep Learning (2020 Fall)** class under Prof. Sang Hyun Park.

Course Description
- Course covered topics in Deep Learning for Image Segmentation, Detection, and Classification. I facilitated office hours, lab sessions, material development, and exam design for 50+ students.

Teaching & Supervision (continued)

Responsibilities
- Grading assignments/exams and providing feedback to students.
- Delivering mini-lectures on specific topics.
- Helping students in the semester project.

Gradute Course
- TA for **Machine Learning (2019 Spring)** class under Prof. Sang Hyun Park.

Course Description
- Course covered topics in machine learning for classic machine learning algorithms, such as SVM, HMM and KNN etc. I facilitated office hours, lab sessions, material development, and exam design for 40+ students.

Responsibilities
- Grading assignments/exams and providing feedback to students.
- Delivering mini-lectures on specific topics.
- Conducting lab experiments to help student.

Supervision

Undergrad
- Supervised undergrad/intern students on final-year projects.

Graduate
- Provide supervision to Master/PhD students on computer vision and deep learning projects.

Projects Participation

2022 – in progress
- *Drone Tracking in IR Videos with Transformers for Anti-Drone System.*
- *Video Domain Adaptation for Semantic Segmentation.*
- *Protecting Personal Computers from Unauthorized Mobile Recordings.*

2018 – 2022
- *Surgical robots segmentation and tracking in X-ray sequences based on deep learning.*
- *Real-Time Tracking of Guidewire Robot Tips Using Deep CNN.*
- *Development and commercialization of micro/nano robot system for precise treatment of brain diseases and tumors.*

2015 – 2017
- *Vehicles Information Extraction Based on Deep Neural Network.*
- *Face Recognition System Based on Deep Learning in Embedded System (Jetson TX1).*
- *Deep Learning Based License Plate Detection.*

Skills

Languages
- Strong reading, writing, and speaking competencies in English. IELTS score 6/9 and TOEIC (710/990).

Coding
- Python, C++.

Deep Learning Libraries
- Tensorflow, Keras, PyTorch, Caffe.

Image Processing Libraries
- OpenCV, DLib C++.

Robotics Libraries
- Robot Navigation and Control using ROS.

Web Dev
- HTML, CSS, Apache Web Server.

Misc.
- Academic research, teaching, training, consultation, LaTeX typesetting and publishing.

Miscellaneous Experience

Presentations at Conferences, Symposia , Workshops, Seminar and Webinars

2022
- Presented a webinar on Catheter and Guidewire Segmentation via Adversarial Learning. in Pak-Austria Fachhochschule: Institute of Applied Sciences and Technology, Pakistan.

2019
- Poster presentation in IPIU Conference, Jeju, South Korea.

2018
- Poster presentation in Conference of KIISE,Jeju, South Korea.

2017
- Oral presentation in Information and Control Symposium, ICS'2017, Jeonju, South Korea.

2016
- Oral presentation in Conference of KISM and SEBS,April29-30, 2016, Busan, South Korea.

Miscellaneous Experience (continued)

Awards and Achievements

- 2024 **DAAD AInet Fellowship**, Selected as a DAAD AInet Fellow (2024).
- 2018-2022 **Scholarship Award**, DGIST Fully Funded Scholarship for PhD.
- 2015-2017 **Scholarship Award**, Chonbuk National University, Fully Funded Scholarship for M.S.
- **Scholarship Award**, Chonbuk National University, Monthly Stipend from Brain Korea 21 Plus (BK-21 Plus).

Summer Schools

- 2019 **DGIST Summer School for Machine learning**. Awarded by DGIST.
- 2018 **DGIST Summer School for Medical Robot**. Awarded by DGIST.

Institutional and Professional Activities

- Co-founder and member of DISA in DGIST.
- Regional Coordinator Pakistani Student Association (PSAK) Korea in DGIST.
- Active member of Pakistani Student Association (PSAK).

References

Prof. Sang Hyun Park, Ph.D
Associate Professor
Robotics and Mechatronics Engineering,
DGIST,
Daegu, South Korea.
shpark13135@dgist.ac.kr

Prof. Hyunki Lee, Ph.D
Professor
Interdisciplinary Engineering,
DGIST,
Daegu, South Korea.
hklee@dgist.ac.kr

Dr. Philip Chikontwe, Ph.D
Postdoctoral Research Fellow
Department of Biomedical Informatics,
Harvard Medical School, USA.
philip_chikontwe@hms.harvard.edu

Research Statement

Research Contributions

My research focuses on advancing deep learning methods for medical image segmentation and tracking, with a particular emphasis on applications in robot-assisted intervention. While deep learning has revolutionized computer vision, including medical imaging tasks such as object segmentation and pixel-level tracking, its success hinges on access to large annotated datasets—a significant challenge in the medical domain. For instance, in robot-assisted interventions, tasks like catheter segmentation are critical for providing visual and haptic feedback to surgeons. Precise tracking of the catheter tip in real-time is necessary to control the robot's position and pose, but the small size and thin shape of the tip make this task highly challenging. Additionally, clinical practice demands delineation of not just the catheter tip but the entire catheter and guidewire to minimize potential vessel damage caused by contact.

In my research, I focus on addressing critical challenges in medical imaging by leveraging synthetic data generation, domain adaptation techniques, and designing advanced deep learning models for improved segmentation. My key contributions include:

- **Efficient Catheter Tracking:** Designing advanced methods to accurately track catheter tips in phantom data, facilitating the development of robot-assisted interventions.
- **Realistic Image Synthesis:** Creating realistic synthetic images and employing them for segmenting catheters and guidewires in real X-ray angiographic sequences.
- **Domain Adaptation Frameworks:** Proposing novel frameworks to minimize domain shifts, enabling models trained on synthetic data to perform effectively in real-world applications.
- **Anatomical Structure Segmentation:** Developing deep learning frameworks for precise segmentation of anatomical structures in medical imaging.

Guidewire & Catheter Tracking in Phantom

Accurate tracking of catheter and guidewire tips is critical for robot-assisted cardiac surgeries. Due to their small size and rapid movement, real-time tracking is challenging. To address this, I developed a deep learning-based tracking method combining detection and segmentation. The system uses a detection network for rough localization and a segmentation network for precise tip delineation. By learning spatial and motion features from successive frames, it improves robustness. During inference, the detection network estimates the initial bounding box, while the segmentation network refines tracking in subsequent frames. If tracking fails, the detection network reinitializes it. This method achieves a real-time tracking speed of 19ms, offering superior accuracy and robustness over existing methods. It provides a reliable solution for stabilizing robot-assisted surgeries. Results from different stages of this research have been published in IEEE conference [1] and IEEE ACCESS Journal [2].

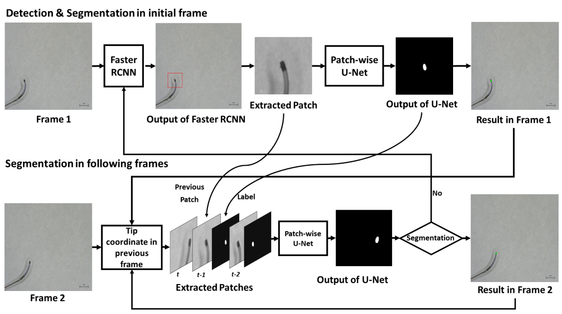

Figure : Overview of the proposed tracking method.

Fluoroscopic Catheter Generation using Generative Modeling

Catheter and guidewire segmentation is essential for robot-assisted interventions guided by fluoroscopy but is often challenged by limited annotated datasets and the complexity of data collection. Deep learning methods, in particular, require large amounts of labeled data, which restricts their applicability in medical domains with scarce resources. To overcome these challenges, I developed a synthesize and segment approach that combines image synthesis and segmentation within a semi-supervised framework. The method leverages an adversarially trained image-to-image translation network to generate realistic synthetic X-ray images containing catheters, effectively augmenting the dataset. To ensure the synthetic images closely resemble real data, the translation network incorporates perceptual loss and similarity constraints, improving the quality and realism of the generated images. These synthesized images are then used with existing segmentation networks to localize catheters accurately, even when labeled samples are limited. Empirical results on medical datasets demonstrate significant improvements over traditional methods, highlighting the potential of this approach to advance catheter segmentation in low-data scenarios. Results from different stages of this research have been published in MICCAI-PRIME [3] and Applied Sciences Journal [4].

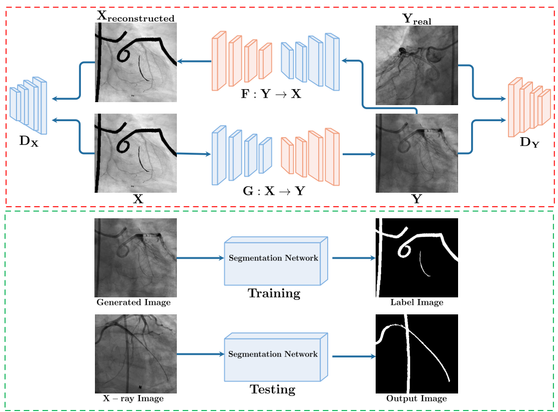

Figure : Framework of the proposed synthesis and segmentation of catheters in X-ray.

Domain Adaptive Segmentation for Fluoroscopic Catheters via Generative Modeling (PhD Thesis)

In this work, we presented an unsupervised domain adaptation (UDA) for catheter segmentation in fluoroscopic sequences using perceptual consistency matching (PCM). Since pixels in two successive frames of a fluoroscopic sequence are perceptually similar and have similar visual features, we assume that a generator (segmentation network) should allocate similar pixels for successive frames of the same class (Catheter or BG). Based on the perceptual feature maps of successive frames, we determine how closely segmentation matches cross-frame pixel correspondence. In the PCM module, we determine pixels with highest correlation across neighboring frames using cosine similarity, and then align the segmentation maps obtained from the generator decoder. If the segmentation matches with PCM-acquired perceptual correlations, then we regard pixels in neighboring frames as belonging to the same class. Note that both intra-domain and inter-domain frame-to-frame PCM is considered in our method. We showed the performance of the proposed method on the Generated-Catheter→Real-Catheter dataset. The results of this research have been published in my PhD thesis.

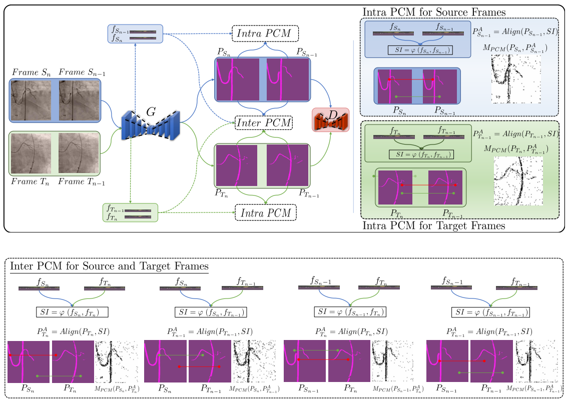

Figure : The proposed domain adaptive segmentation of catheters in X-ray.

Video Domain Adaptive Segmentation via Generative Modeling in Computer Vision

UDA is key for transferring knowledge from labeled datasets (sources) to improve performance on new unlabeled datasets (targets). While image-based UDA has seen significant advancements, video-based UDA remains less explored due to challenges in adapting diverse video features and modeling temporal associations. Existing methods often rely on optical flow to capture motion cues between frames, but this approach is computationally expensive and struggles with domain diversity. To overcome these challenges, I developed an adversarial domain adaptation approach for video semantic segmentation that eliminates the need for optical flow. At the core of this approach is the PCM strategy, which uses perceptual similarity to align correlated pixels across consecutive frames, ensuring they correspond to the same class. By enforcing consistency within and across domains, this method significantly improves prediction accuracy. Extensive experiments on public datasets demonstrate that my approach outperforms existing UDA methods and offers improved efficiency with faster inference times. Results from this research have been published in the Neural Network Journal [5].

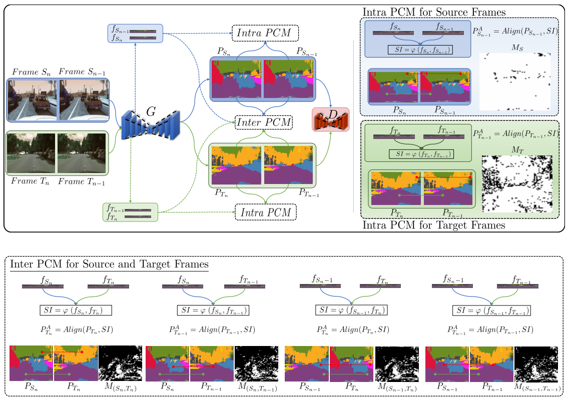

Figure : The proposed video domain adaptation framework for semantic segmentation.

Anatomical Structure Segmentation in Chest X-ray

Automated multi-organ segmentation is crucial for computer-aided diagnostic (CAD) systems in chest X-ray fluoroscopy. However, indistinct structures, individual anatomical variations, and artifacts like pacemakers and catheters make accurate segmentation difficult. To address these issues, I developed a robust deep-learning framework for segmenting anatomical structures in chest radiographs using a dual encoder-decoder convolutional neural network (CNN). The first network in the framework employs a pre-trained VGG19 encoder, enhanced with a Squeeze-and-Excitation (SE) module to improve feature representation, and a decoder to generate an initial segmentation mask. This mask is combined with the input image and passed through a second encoder-decoder network with recurrent residual blocks and an attention gate module to capture additional contextual details and refine smaller regions. The results show that the proposed method outperformed existing techniques, demonstrating its effectiveness for both multi-class and single-class segmentation tasks. This research has been published in Scientific Reports [6].

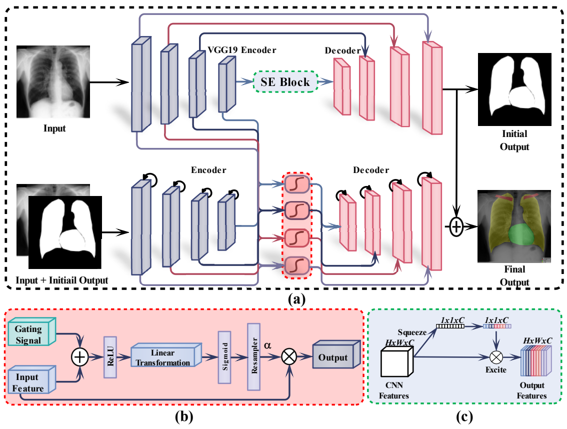

Figure : (a) The proposed segmentation framework, (b) AGM with attention coeff α scaling features, and (c) SE block.

Future Research Plans

My previous research tackles critical challenges in medical imaging and video-based domain adaptation, with a focus on improving segmentation tasks such as catheter localization in X-ray fluoroscopy through synthetic data generation and adversarial domain adaptation. Building on this work, my future research aims to advance video domain adaptation and enhance robustness for real-world medical applications. Specifically, I plan to develop efficient and generalizable methods for video semantic segmentation and image-based tasks that perform well across diverse domains with minimal labeled data. To achieve this, I will integrate state-of-the-art models like Stable Diffusion [7] and Vision Language Models (VLMs) [8]. A key aspect of my work will be improving synthetic data generation for medical imaging tasks, such as catheter and guidewire segmentation, by leveraging Stable Diffusion to create realistic, motion-aware fluoroscopic images that enhance model robustness, especially in low-data scenarios. Furthermore, I plan to utilize VLMs to enhance video-based domain adaptation, combining visual, motion, and textual features for more efficient, context-aware scene understanding. Motivated by my expertise, I envision working with a research team focused on developing reliable and robust deep-learning models by addressing the following challenges.

Data Generation using Generative Modeling

Building on our work in synthesizing realistic catheter images for data augmentation in fluoroscopy[3,4], we aim to explore the integration of Stable Diffusion models for more realistic and controllable image generation. Stable Diffusion, with its capacity for high-quality, diverse, and detailed image synthesis, can potentially bridge the gap between synthetic and real-world fluoroscopic images. This will enhance the robustness of downstream segmentation tasks, particularly in low-data regimes. Moreover, we propose leveraging motion cues inherent in fluoroscopic sequences to generate temporally coherent synthetic data. By incorporating motion information, we can create dynamic datasets that better mimic real-world conditions encountered in robot-assisted interventions. These advancements aim to:

- **Enhance data diversity:** Use motion-aware synthesis to simulate complex scenarios of catheter and guidewire manipulation.
- **Improve segmentation performance:** Train models on temporally consistent, realistic data to improve generalizability across different patient anatomies and imaging conditions.
- **Facilitate domain adaptation:** Utilize motion-guided Stable Diffusion to generate datasets tailored to varying clinical setups and imaging equipment.

Video Domain Adaptation with VLM and Stable Diffusion for Efficient and Context-Aware Scene Understanding

Following my research 'video-based unsupervised domain adaptation (video UDA)' [5], I aim to integrate Stable Diffusion and VLMs for data generation and model training. Stable Diffusion will be utilized to generate synthetic data with realistic spatial and temporal dynamics, addressing the challenge of domain variability in video segmentation tasks. Subsequently, these generated datasets will serve as the foundation for training VLMs to enable richer scene understanding by combining visual, motion, and textual features.

I am particularly interested in integrating temporal and textual information, as demonstrated in approaches like FairCLIP [9], ZegCLIP [10], and cross-class domain adaptive segmentation with VLM [11]. While optical flow models are traditionally employed for motion understanding, their computational intensity makes them impractical for real-time medical imaging. To address this, we aim to leverage VLMs for efficient extraction and utilization of temporal features, offering the following advancements:

- **Efficient motion representation:** Use VLMs to capture temporal dynamics without relying on optical flow, enabling real-time processing in resource-constrained environments.
- **Multimodal integration:** Fuse visual features, motion cues, and textual descriptions to enhance scene understanding, particularly in medical imaging scenarios where contextual insights are critical.
- **Domain adaptation robustness:** Utilize generated data and VLMs to bridge the domain gap between synthetic and real-world datasets, improving the model's adaptability and segmentation accuracy.

Segmentation Models with Attention

Building on my previous work in multi-organ segmentation in chest X-ray fluoroscopy [6], I aim to explore advanced segmentation models with attention mechanisms, particularly leveraging transformer-based architectures. Transformers have revolutionized computer vision, especially for tasks requiring fine-grained understanding, by capturing long-range dependencies and modeling global contextual information. My goal is to integrate transformers with convolutional neural networks (CNNs) to create hybrid models that combine the localized feature extraction of CNNs with the global context awareness of transformers. Specifically, I plan to develop transformer-based segmentation models tailored for medical imaging, such as multi-organ segmentation in fluoroscopy and CT scans. By incorporating self-attention mechanisms and leveraging architectures like the Vision Transformer (ViT) [12] and Swin Transformer [13], I aim to address challenges like indistinct boundaries, anatomical variability, and the presence of artifacts. This research will also investigate attention-guided mechanisms for refining segmentation in regions with complex structures or small anatomical details.

References:

[1] Ullah, I., Chikontwe, P. and Park, S.H., *Guidewire tip tracking using U-Net with shape and motion constraints*, 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pp. 215–217, IEEE, 2019.

[2] Ullah, I., Chikontwe, P. and Park, S.H., *Real-time tracking of guidewire robot tips using deep convolutional neural networks on successive localized frames*, IEEE Access, 7, pp. 159743–159753, IEEE, 2019.

[3] Ullah, I., Chikontwe, P. and Park, S.H., *Catheter synthesis in X-ray fluoroscopy with generative adversarial networks*, In Predictive Intelligence in Medicine: Second International Workshop, PRIME 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 2, Springer International Publishing, 2019.

[4] Ullah, I., Chikontwe, P., Choi, H., Yoon, C.H. and Park, S.H., *Synthesize and segment: Towards improved catheter segmentation via adversarial augmentation*, Applied Sciences, 11(4), p.1638, 2021.

[5] Ullah, I., An, S., Kang, M., Chikontwe, P., Lee, H., Choi, J. and Park, S.H., *Video domain adaptation for semantic segmentation using perceptual consistency matching*, Neural Networks, 179, p.106505, 2024.

[6] Ullah, I., Ali, F., Shah, B., El-Sappagh, S., Abuhmed, T. and Park, S.H., *A deep learning based dual encoder–decoder framework for anatomical structure segmentation in chest X-ray images*, Scientific Reports, 13(1), p.791, Nature Publishing Group UK London, 2023.

[7] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B., *High-Resolution Image Synthesis with Latent Diffusion Models*, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, 10684-10695. https://doi.org/10.1109/CVPR52688.2022.01044, 2022.

[8] Radford, A., Kim, J.W., Hallacy, C., et al., *Learning Transferable Visual Models From Natural Language Supervision*, Proceedings of the International Conference on Machine Learning (ICML),8748-8763., 2021.

[9] Luo, Y., Shi, M., Khan, M.O., Afzal, M.M., Huang, H., Yuan, S., Tian, Y., Song, L., Kouhana, A., Elze, T. and Fang, Y., *Fairclip: Harnessing fairness in vision-language learning*, In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12289–12301, 2024.

[10] Zhou, Z., Lei, Y., Zhang, B., Liu, L. and Liu, Y., *Zegclip: Towards adapting clip for zero-shot semantic segmentation*, In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11175–11185, 2023.

[11] Ren, W., Xia, R., Zheng, M., Wu, Z., Tang, Y. and Sebe, N., *Cross-Class Domain Adaptive Semantic Segmentation with Visual Language Models*, In Proceedings of the 32nd ACM International Conference on Multimedia, pp. 5005–5014, October, 2024.

[12] Dosovitskiy, A., *An image is worth 16x16 words: Transformers for image recognition at scale*, arXiv preprint arXiv:2010.11929, 2020.

[13] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., and Wang, Y., *Swin Transformer: Hierarchical Vision Transformer using Shifted Windows*, In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 10012–10022, 2021.

Teaching Statement

As an academic deeply passionate about advancing knowledge in Artificial Intelligence (AI), deep learning, computer vision, and machine learning, I find immense fulfillment in both teaching and mentoring. My approach centers on fostering an engaging, inclusive, and dynamic learning environment where students can connect theoretical concepts with real-world applications.

Teaching Philosophy
My teaching philosophy revolves around three core principles:

- **Bridging Theory and Practice:** I aim to connect foundational theories with hands-on experimentation, enabling students to grasp both the 'how' and the 'why' behind AI methodologies. Through coding exercises, collaborative projects, and problem-solving tasks, I prepare students for practical challenges in academia and industry.
- **Research-Driven Learning:** I integrate cutting-edge research into the curriculum, allowing students to participate in projects addressing real-world challenges. This not only equips them with technical skills but also nurtures critical thinking and innovation.
- **Inclusive Mentorship:** Recognizing diverse learning needs, I strive to provide personalized guidance to ensure all students feel empowered to achieve their goals. By creating an open and collaborative environment, I encourage curiosity and teamwork.

Teaching Experience
I bring a blend of teaching and mentoring experience at both undergraduate and graduate levels. At DGIST, I designed and taught *Applied AI for Intelligent Robotic Systems (IE803, Spring 2024)*, a course focused on AI-driven applications. It bridged foundational topics in neural networks with advanced concepts such as transformers, while incorporating hands-on lab sessions and project-based evaluations to equip students with practical AI skills.

Building on this foundation, I am currently developing a new course, *Intelligent Robots (IE916, Spring 2025)*, which shifts the focus toward intelligent robotics by integrating AI concepts with the Robot Operating System (ROS). This course emphasizes the design and implementation of intelligent robotic systems, blending theoretical understanding with practical applications in robotics, navigation, and automation. It aims to prepare students for cutting-edge research and industry roles in robotics and AI.

Graduate-Level Courses Taught:

- *Applied AI for Intelligent Robotic Systems (IE803, Spring 2024)* |Course-Material| |Codes| |Course Evaluation|
  - This course provided a comprehensive exploration of Applied Artificial Intelligence (AI) in the context of intelligent robotic systems. Through theoretical lectures, practical exercises, and hands-on projects, students gained a thorough understanding of the integration of AI approaches in the development and functioning of robotic systems. The course curriculum covered a variety of topics, including machine learning approaches, computer vision, and other artificial intelligence techniques specifically tailored for robotics. More specifically, the course explored applications such as classification, segmentation, recognition, detection, and tracking. Students applied AI-driven solutions to real-world scenarios, developing practical expertise essential for modern robotics engineering. Project-based work allowed students to implement AI-driven solutions in their research endeavors.
- *Intelligent Robotic Systems (IE803, Spring 2025, planned)*
  - This course is designed to offer a comprehensive exploration of robotics, focusing on foundational principles and practical applications in robotic systems. Students will learn key robotics concepts, including kinematics, dynamics, and control, through hands-on activities such as solving kinematics problems, tuning PID controllers, and generating robot trajectories. A significant emphasis is placed on the Robot Operating System (ROS), where students will progress from basic commands and TurtleBot3 simulations to advanced motion planning and trajectory execution with the UR10 robotic arm. The course will also integrate machine learning and convolutional neural networks (CNNs) to enable real-world applications like classification, detection, and tracking, enhancing the robotics systems' adaptability and intelligence.

Additionally, I have served as a teaching assistant for:

- *Deep Learning (2020 Fall):* Facilitated lab sessions, graded assignments, and guided students on projects in image segmentation, detection, and classification. |Course-Info|
- *Machine Learning (2019 Spring):* Assisted students with concepts such as SVM, HMM, and KNN through lab experiments and feedback on coursework. |Course-Info|

Courses I Can Teach

Undergraduate Courses	**Graduate Courses**
Artificial Intelligence	Computer Vision
Machine Learning	Advanced Image Processing
Image Processing	Advanced Topics in Deep Learning
Deep Learning	Medical Image Analysis
Introduction to Programming	Applied AI for Intelligent Robotic Systems
Object-Oriented Programming	Intelligent Robots

Supervision and Mentorship

I have actively mentored students at various academic levels, fostering their growth through:

- **Graduate Research:** Guided Master's and Ph.D. students in exploring advanced topics such as semantic segmentation and robustness in AI models. My students have achieved recognition in challenges like the *STIR-MICCAI Challenge (2024)* and *MICCAI-SegStrong Challenge (2024)*.
- **Undergraduate Projects:** Supervised final-year projects and internships focused on computer vision and deep learning.
- **Undergraduate Research Program (UGRP):** Supervised undergraduate students in the development of a robust deep learning approach for the segmentation of surgical tools in videos. This collaborative effort culminated in the publication of our findings in the 2018 Robotic Scene Segmentation Challenge.

Future Directions in Teaching

Looking ahead, I aspire to expand my teaching portfolio by incorporating:

- **Interdisciplinary Courses:** Developing courses that merge AI with domains like healthcare, robotics, and sustainability.
- **Project-Based Learning:** Introducing capstone projects that mirror real-world challenges, enhancing students' readiness for industry and research.
- **Emerging Topics:** Integrating advancements such as Vision Language Models (VLMs) and generative AI into curricula.

Through teaching, I aim to inspire the next generation of AI practitioners and researchers, equipping them with the skills and mindset to address tomorrow's challenges.

Teaching References

Prof. Sang Hyun Park, Ph.D
Associate Professor
Robotics and Mechatronics Engineering,
DGIST, Daegu, South Korea.
shpark13135@dgist.ac.kr

Prof. Hyunki Lee, Ph.D
Professor
Interdisciplinary Engineering,
DGIST, Daegu, South Korea.
hklee@dgist.ac.kr