

RAFI IBN SULTAN

PhD Candidate in Computer Science, Wayne State University
rafi.ruet13@gmail.com | [LinkedIn](#) | [GitHub](#) | [Website](#) | [Google Scholar](#)

RESEARCH PROFILE

My research centers on grounded multimodal intelligence, building vision-language systems whose reasoning and predictions remain anchored to visual evidence for reliable use in complex real-world settings.

EDUCATION

PhD Computer Science *Expected August, 2027*

Advisor: Prof. Dongxiao Zhu

Wayne State University, Detroit, MI, USA

MSc Computer Science *December 2025*

Wayne State University, Detroit, MI, USA

B.Sc. Computer Science and Engineering *November 2018*

Rajshahi University of Engineering & Technology (RUET), Bangladesh

RESEARCH EXPERIENCE

Graduate Research Assistant Wayne State University *September 2023 – Present*
Detroit, MI, USA

- Lead end-to-end research projects in multimodal AI, including problem formulation, model development, experimental design, evaluation, and preparation of research artifacts and publications.
- Lead an ongoing clinical collaboration with Henry Ford Health on coronary artery segmentation from non-contrast planning CT, working with clinical collaborators to translate domain requirements into model design and clinically meaningful evaluation.
- Mentor junior PhD students on research methodology, experimental design, implementation, and academic writing.

TEACHING AND MENTORING

Graduate Teaching Assistant Wayne State University *August 2022 – August 2023*
Detroit, MI, USA

- Led Software Engineering labs, supervising 5 to 6 project teams of 4 to 5 students each per term through the full software development lifecycle, from requirements to delivery.

Lecturer Varendra University *October 2019 – August 2022*
Rajshahi, Bangladesh

- Served as instructor of record for undergraduate courses including Microprocessor and Assembly Language, Introduction to Object-Oriented Programming, and Computer Fundamentals, with full responsibility for lectures, assessments, and grading.
- Supervised approximately 8 undergraduate final-year thesis and capstone projects from proposal to defense.

First-Author Publications

- [1] **RI Sultan**, H Zhu, X Zhou, C Li, P Khanduri, M Brocanelli, D Zhu. “WalkGPT: Grounded Vision-Language Conversation with Depth-Aware Segmentation for Pedestrian Navigation,” *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026.
- [2] **RI Sultan**, C Li, Y Demetriou, AI Ghanem, JP Kim, J Cunningham, H Bagher-Ebadian, D Zhu, KS Thind. “A Neighborhood Attention Transformer Network for Enhanced 3D Segmentation of the Left Anterior Descending Artery,” *Medical Physics*, 2026.
- [3] **RI Sultan**, H Zhu, C Li, D Zhu. “MedPlex: Deep Vision-Language Co-Adaptation for Clinically Grounded Medical Segmentation,” *British Machine Vision Conference (BMVC)*, 2026.
- [4] **RI Sultan**, C Li, H Zhu, P Khanduri, M Brocanelli, D Zhu. “GeoSAM: Fine-Tuning SAM with Multi-Modal Prompts for Mobility Infrastructure Segmentation,” *European Conference on Artificial Intelligence (ECAI)*, 2025.

First-Author Manuscripts (Under Review and Preprints)

- [5] **RI Sultan**, X Zhou, MSA Chowdhury, C Li, P Khanduri, M Brocanelli, D Zhu. “Spatial-CORE: Spatially Confident Reasoning for Large Vision-Language Models,” under review, 2026.

Collaborative Publications

- [6] P Khanduri, C Li, **RI Sultan**, A Sarker, Y Qiang, J Kliewer, D Zhu. “On Federated Compositional Optimization: Algorithms, Analysis, and Guarantees,” *Transactions on Machine Learning Research (TMLR)*, 2026.
- [7] C Li, **RI Sultan**, H Bagher-Ebadian, Y Qiang, K Thind, D Zhu, IJ Chetty. “Enhancing CT Image Segmentation Accuracy Through Ensemble Loss Function Optimization,” *Medical Physics*, 2025.
- [8] U Mgbob, **RI Sultan**, J Kim, K Thind, D Zhu. “FluenceFormer: Transformer-Driven Multi-Beam Fluence Map Regression for Radiotherapy Planning,” *Medical Imaging with Deep Learning (MIDL)*, 2026.
- [9] C Li, **RI Sultan**, P Khanduri, Y Qiang, C Indrin, D Zhu. “AutoProSAM: Automated Prompting SAM for 3D Multi-Organ Segmentation,” *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2025.
- [10] C Li, H Zhu, **RI Sultan**, HB Ebadian, P Khanduri, C Indrin, K Thind, D Zhu. “MulModSeg: Enhancing Unpaired Multi-Modal Medical Image Segmentation with Modality-Conditioned Text Embedding and Alternating Training,” *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2025.
- [11] C Li, H Bagher-Ebadian, **RI Sultan**, M Elshaikh, B Movsas, D Zhu, IJ Chetty. “A New Architecture Combining Convolutional and Transformer-Based Networks for Automatic 3D Multi-Organ Segmentation on CT Images,” *Medical Physics*, 2023.
- [12] C Li, Y Qiang, **RI Sultan**, H Bagher-Ebadian, P Khanduri, IJ Chetty, D Zhu. “FocalUNETR: A Focal Transformer for Boundary-Aware Prostate Segmentation Using CT Images,” *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2023.

INVITED TALKS

- “WalkGPT: Pixel-Grounded Navigation Guidance for Pedestrians,” **Best of CVPR 2026 Series, Voxel51**, July 2026 (virtual).
- “WalkGPT: Grounded Vision-Language Conversation with Depth-Aware Segmentation for Pedestrian Navigation,” **Computer Vision Group, Cohere Labs Open Science Community**, May 2026 (virtual).
- “Multimodal AI for Mobility Infrastructure and Pedestrian Guidance,” **Departmental Seminar, Wayne State University**, 2026.

AWARDS AND HONORS

- **Outstanding Graduate Research Assistant Award**, Department of Computer Science, Wayne State University, 2025.

PROFESSIONAL SERVICE

- **Conference Reviewer:** AAAI 2027, NeurIPS 2026; IJCAI-ECAI 2026; BMVC 2026; IJCNN 2025.
- **Journal Reviewer:** Pattern Recognition; Computer Vision and Image Understanding (CVIU); IEEE Transactions on Circuits and Systems for Video Technology (TCSVT); Knowledge-Based Systems; Computer Methods and Programs in Biomedicine; Computers in Biology and Medicine; Biomedical Signal Processing and Control; Results in Engineering.

TECHNICAL EXPERTISE

- **Programming and Systems:** Python, PyTorch, Linux, Git, LaTeX, multi-GPU deep learning workflows.
- **Models and Methods:** Vision transformers, SAM, large vision-language models, LoRA/parameter-efficient adaptation, contrastive learning, segmentation, grounding, representation learning.
- **Tools and Libraries:** Hugging Face, MONAI, OpenCV, scikit-learn, Weights & Biases.

Last updated: August 2026