Kashu Yamazaki

kyamazak@andrew.cmu.edu
kashu7100.github.io
Scholar ID: TF2LRvMAAAAJ

Education

Ph.D., Language and Information Technology, **Carnegie Mellon University**, Pittsburgh, PA — On going
Advisor: Dr. Fragkiadaki, Katerina

M.S., Computer Science and Computer Engineering, **University of Arkansas**, Fayetteville, AR — Aug 2023
Thesis: Towards Multi-Modal Explainable Video Understanding — Advisor: Dr. Le, Ngan

- Cumulative GPA: 3.875/4

B.S., Mechanical Engineering, *Summa Cum Laude*, **University of Arkansas**, Fayetteville, AR — Dec 2020
Thesis: Towards Sensorimotor Coupling of a Spiking Neural Network and Deep Reinforcement Learning for Robotics Application — Advisor: Dr. Zhou, Wenchao

- Minor in Computer Science
- Major GPA: 4.0/4, Minor GPA: 4.0/4, Cumulative GPA: 3.952/4

Research Experience

CMU Katef group, PhD Student, Pittsburgh, PA — Sep 2024 - Present
Advised by Dr. Fragkiadaki, Katerina

- Robotics: Generative simulation for robot learning [2]

CMU MLSP group, Visiting Researcher, Pittsburgh, PA — Nov 2023 - Apr 2024
Advised by Dr. Ramakrishnan, Bhiksha Raj

- Vision-Language: Robustness of Referring Perception [4]

UA AICV Lab, **Research Associate**, Fayetteville, AR — Sep 2023 - Jul 2024
PI: Dr. Le, Ngan

- Robotics: Queryable Mapping [6], Non-Rigid 3D Reconstruction [8]

UA AICV Lab, **Research Assistant**, Fayetteville, AR — Jan 2020 - Aug 2023
Advised by Dr. Le, Ngan

- Vision-Language: Interpretable video representation [3] [9] [11] [13]
- Video Understanding: TAPG [11] [14] [16] [19], VPC [9] [13], and Anomaly Detection [7]
- Medical: Segmentation [17] [18] [22] [21], NAS [25]

UA MSR Lab, **Undergraduate Research Assistant**, Fayetteville, AR — Dec 2018 - Dec 2019
Advised by Dr. Chen, Yue (now at Georgia Tech)

- Soft robot development for photodynamic therapy of pancreatic cancer [20] and ovarian cancer [23]
- Soft robotic gripper development for delicate object grasping

Industry Experience

Deloitte Consulting, *Consultant Intern (Remote)*, — Aug 2022

- Proposed a HR strategy for an insurance company with a special focus on digital transformation (DX).

TeirIV, Inc., *ML Engineering Intern (Remote)*, — Jun 2021 - Oct 2022

- Proposed an unsupervised model evaluation protocol for semantic segmentation models based on the ELBO of a categorical diffusion model for internal model evaluation.
- Implemented PointPillars with aleatoric uncertainty estimation for 3D object detection with point clouds.

Teaching Experience

NACME-Google, **Applied Machine Learning Intensive**, Fayetteville, AR — Su 2021
UA, **MEEG 2003** (Statics), Fayetteville, AR — Fa 2017 - Fa 2018

Professional Service

Conference Reviewer: CVPR (24, 25), ICCV (25), AAAI (23 - 25), ICASSP (23, 24), ICML (22), ICIP (22)
Journal Reviewer: TPAMI (Since 2024), TNNLS (Since 2023), ISA Transactions (Since 2023)
Membership: *Tau Beta Pi*, the Engineering Honor Society (Since 2017)

Patents

[1] M. Bai, Y. Chen, Y. Liu, Y. Li, and **K. Yamazaki**, Soft Robotic Laparoscope for Minimally Invasive Intraperitoneal Photodynamic Therapy, USPTO, U.S. Provisional Patent Application No. 62/967825, 2020.

Software

[2] Genesis Authors "Genesis: A Universal and Generative Physics Engine for Robotics and Beyond". 2024.

Peer-Reviewed Publications

[3] K. Vo, T. Phan, **K. Yamazaki**, M. Tran, N. Le, "HENASY: Learning to Assemble Scene-Entities for Interpretable Egocentric Video-Language Model". NeurIPS (**Poster**). 2024.

[4] X. Li, K. Qiu, J. Wang, X. Xu, R. Singh, **K. Yamazaki**, H. Chen, X. Huang, B. Raj, "R^2-Bench: Benchmarking the Robustness of Referring Perception Models under Perturbations". ECCV (**Poster**), 2024.

[5] T. Hanyu*, **K. Yamazaki***, M. Tran, R.A. McCann, H. Liao, C. Rainwater, M. Adkins, J. Cothren and N. Le, "AerialFormer: Multi-Resolution Transformer for Aerial Image Segmentation". Remote Sensing, 16(16), p.2930. 2024

[6] **K. Yamazaki**, T. Hanyu, K. Vo, T. Pham, M. Tran, G. Doretto, A. Nguyen, N. Le "Open-Fusion: Real-time Open-Vocabulary 3D Mapping and Queryable Scene Representation," ICRA (**Oral**), 2023.

[7] H. Joo, K. Vo, **K. Yamazaki**, N. Le "CLIP-TSA: CLIP-Assisted Temporal Self-Attention for Weakly-Supervised Video Anomaly Detection," ICIP (**Oral**), 2023.

[8] K. Vo, T. Pham, **K. Yamazaki**, M. Tran, N. Le "DNA: Deformable Neural Articulations Network for Template-Free Dynamic 3D Human Reconstruction From Monocular RGB-D Video," CVPRW, 2023.

[9] **K. Yamazaki**, K. Vo, S. Truong, B. Raj, N. Le "VLTinT: Visual-Linguistic Transformer-in-Transformer for Coherent Video Paragraph Captioning," AAAI (**Oral**), 2023.

[10] M. Tran, K. Vo, **K. Yamazaki**, A. Fernandes, M. Kidd, N. Le "AISFormer: Amodal Instance Segmentation with Transformer," BMVC, 2022.

[11] K. Vo, S. Truong*, **K. Yamazaki***, B. Raj, M. Tran, N. Le "AOE-Net: Entities Interactions Modeling with Adaptive Attention Mechanism for Temporal Action Proposals Generation," IJCV (**IF: 19.5**), 2022.

[12] **K. Yamazaki**, K. Vo, D. Bulsara, N. Le "Spiking Neural Networks and Their Applications: A Review," Brain Sciences (**Editor's Choice**; **Best Paper Award**), 2022.

[13] **K. Yamazaki**, S. Truong, K. Vo, M. Kidd, C. Rainwater, K. Luu, N. Le "VLCap: Vision-Language with Contrastive Learning for Coherent Video Paragraph Captioning," ICIP (**Oral**), 2022.

[14] K. Vo, H. Joo*, **K. Yamazaki***, S. Truong, K. Kitani,, M.-T. Tran, N. Le "AEI: Actors-Environment Interaction with Adaptive Attention for Temporal Action Proposals Generation," BMVC (**Oral-3.33%**), 2021.

[15] N. Le, V. Rathour*[1], **K. Yamazaki***[1], K. Luu, and M. Savvides "Deep Reinforcement Learning in Computer Vision: A Comprehensive Survey," Artificial Intelligence Review (**IF: 12.0**), 2021.

[16] K. Vo, **K. Yamazaki**, S. Truong, M.-T. Tran, A. Sugimoto, and N. Le "ABN: Agent-Aware Boundary Networks for Temporal Action Proposal Generation," IEEE Access, 2021.

[17] N. Le, T. Bui, K. Vo-Ho, **K. Yamazaki**, K. Luu "Narrow Band Active Contour Attention Model for Medical Segmentation," Diagnostics, 2021.

[18] **K. Yamazaki**, N. Le, V. Rathour "Invertible Residual Network with Regularization for Effective Volumetric Segmentation," SPIE Medical Imaging, 2021.

[19] V. Vo-Ho, N. Le, **K. Yamazaki**, A. Sugimoto, and M. Tran "Agent-Environment Network for Temporal Action Proposal Generation," ICASSP, 2021

[20] Y. Li, Y. Liu, **K. Yamazaki**, M. Bai and Y. Chen, "Development of a Soft Robot-Based Photodynamic Therapy for Pancreatic Cancer," in IEEE Transactions on Mechatronics (**IF: 6.4**), 2021.

[21] N. Le, **K. Yamazaki**, K. Quach, D. Truong, and M. Savvides "A Multi-task Contextual Atrous Residual Network for Brain Tumor Detection & Segmentation," ICPR, 2020.

[22] N. Le, T. Le, **K. Yamazaki**, B. Toan, K. Luu "Offset Curves Loss for Imbalanced Problem in Medical Segmentation," ICPR, 2020.

[23] Y. Liu, **K. Yamazaki**, D. Zhang, Y. Li, M. Su, Q. Xie, Y. Chen, and M. Bai, "Minimally Invasive Intraperitoneal Photodynamic Therapy Using a New Soft Robot System," SPIE, 2020.

[24] E. Sirotkin, **K. Yamazaki**, and A. Miroshnichenko. "Gearbox Development for an Emergency Brake System of the Wind Turbine," IOP Conference Series: Earth and Environmental Science, 2020.

Book Chapters

[25] K. Vo, **K. Yamazaki**, H. Hoang, M. Tran, N. Le "Neural Architecture Search for Medical Image Applications". Meta-Learning with Medical Imaging and Health Informatics Applications, 2023.