

Shao-Yuan Lo

Email: sylo@jhu.edu
Phone: 1-443-808-7270
<https://shaoyuanlo.github.io>

EDUCATION

Johns Hopkins University (JHU)	Baltimore, MD
Ph.D. in Electrical and Computer Engineering	2019 –2023
– Thesis: Robust Computer Vision Against Adversarial Robustness and Domain Shifts	
– Committee: Vishal M. Patel (advisor), Rama Chellappa, Alan Yuille, Jesus Villalba	
M.S.E. in Electrical and Computer Engineering	2019 –2023
National Chiao Tung University (NCTU)	Hsinchu, Taiwan
M.S. in Electronics Engineering	2017 –2019
– Thesis: Real-Time Semantic Segmentation Networks for Autonomous Driving	
– Committee: Hsueh-Ming Hang (advisor), Wen-Hsiao Peng, Wen-Huang Cheng	
B.S. in EECS Undergraduate Honors Program	2013 –2017
University of Illinois at Urbana-Champaign (UIUC)	Champaign, IL
Exchange Program in Electrical and Computer Engineering	2016

INDUSTRY EXPERIENCE

Honda Research Institute USA	San Jose, CA
Research Scientist	Jul. 2023 –Present
– Cognition Team	
– Managers: Kwonjoon Lee and Behzad Dariush	
Amazon Web Services	Seattle, WA
Applied Scientist Intern	May 2022 –Aug. 2022
– Just Walk Out Technology Team	
– Mentors: Poojan Oza, Sumanth Chennupati, and Alejandro Galindo	
Amazon Lab126	Bellevue, WA
Applied Scientist Intern	May 2021 –Aug. 2021
– Astro Vision Science Team	
– Mentors: Wei Wang, Jim Thomas, Jingjing Zheng, and Cheng-Hao Kuo	

RESEARCH AREAS

My research lies in AI Safety and beyond, focusing on making AI systems safe, as well as leveraging AI to make human society safe. The AI safety problems that I have been contributing to include three aspects: (1) Robustness: ensuring resistance to adverse situations and unforeseen domains (e.g., Adversarial Machine Learning [6-10, 12-13, 15] and Domain Generalization [11, 14]); (2) Monitoring: identifying and forecasting malicious scenarios (e.g., Anomaly Detection [12, 19, 24] and Action Anticipation [17-18] using Multimodal Large Language Models); and (3) Alignment: steering AI to facilitate human-AI harmony (e.g., Affective Reasoning [28], Machine Theory-of-Mind [26], and Efficient Instruction Tuning [25]).

AWARDS

• Robert F. Wagner All-Conference Best Student Paper Award , SPIE Medical Imaging 2024	2024
• CVPR DEI Travel Award , IEEE/CVF CVPR 2023	2023
• Google CS Research Mentorship Program , Google Research	2022
• Government Scholarship to Study Abroad , Ministry of Education, Taiwan	2020
• First-Year Doctoral Fellowship , ECE Department, JHU	2019
• Government Scholarship to World Top 100 Universities (declined), Ministry of Education, Taiwan	2019
• Best Paper Award , ACM Multimedia Asia 2019	2019
• Best Master Thesis Award , Chinese Image Processing and Pattern Recognition Society	2019
• Students' Outstanding Contribution Award (highest honor), NCTU	2019
• Dean's List , EECS Honors Program, NCTU	2017
• Scholarship for Outbound Exchange , NCTU	2016
• WINTEK Outstanding Freshmen Scholarship , WINTEK Corp. and NCTU	2013

PUBLICATIONS

Preprints

- [28] Yuxiang Guo, Faizan Siddiqui, Yang Zhao, Rama Chellappa, and **Shao-Yuan Lo**. “StimuVAR: Spatiotemporal Stimuli-aware Video Affective Reasoning with Multimodal Large Language Models.” In *International Journal of Computer Vision (IJCV)*, 2025. [IF=11.6] (under review, revision stage)
- [27] Jr-Jen Chen, Kuang-Ming Chen, Po-Chun Lu, Chang-Yu Hsu, Aaron Y., Yi-Chin Hsieh, Chieh Tu, Zong-Wei Li, Ting-Kuan Wu, Yu-Chiang Wang, **Shao-Yuan Lo**, and Da-Cheng Juan. “FormCraft: A Three-Level Benchmarking Approach to Form Intelligence.” In *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025. (under review)

Journals and Conferences

- [26] Chunhui Zhang, Sean Dae Houlihan, Kwonjoon Lee, Nakul Agarwal, Zhongyu Ouyang, Soroush Vosoughi, and **Shao-Yuan Lo**. “Overcoming Multi-step Complexity in Multimodal Theory-of-Mind Reasoning: A Scalable Bayesian Planner.” In *International Conference on Machine Learning (ICML)*, 2025. [Spotlight (top 2.6%)]
- [25] Bardia Safaei, Faizan Siddiqui, Jiacong Xu, Vishal M. Patel, and **Shao-Yuan Lo**. “Filter Images First, Generate Instructions Later: Pre-Instruction Data Selection for Visual Instruction Tuning.” In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. [Highlight (top 3%)]
- [24] Jiacong Xu, **Shao-Yuan Lo**, Bardia Safaei, Vishal M. Patel, and Isht Dwivedi. “Towards Zero-Shot Anomaly Detection and Reasoning with Multimodal Large Language Models.” In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. [Highlight (top 3%)]
- [23] Chia-Hao Kao, Cheng Chien, Yu-Jen Tseng, Yi-Hsin Chen, Alessandro Gnutti, **Shao-Yuan Lo**, Wen-Hsiao Peng, and Riccardo Leonardi. “Bridging Compressed Image Latents and Multimodal Large Language Models.” In *International Conference on Learning Representations (ICLR)*, 2025.
- [22] Anirudh Thatipelli, **Shao-Yuan Lo**, and Amit K. Roy-Chowdhury. “Exocentric and Egocentric Methods: A Short Survey.” In *Computer Vision and Image Understanding (CVIU)*, 2025. [IF=4.3]
- [21] Shaoyan Pan, Junbo Peng, Yuan Gao, **Shao-Yuan Lo**, Tianyu Luan, Junyuan Li, Tonghe Wang, Chih-Wei Chang, Zhen Tian, and Xiaofeng Yang. “Volumetric CBCT Reconstruction using Single X-ray Projection Image with Cycle-domain Geometry-integrated Denoising Diffusion Probabilistic Models.” In *IEEE Transactions on Medical Imaging (TMI)*, 2025. [IF=8.9]
- [20] Shaoyan Pan, Vanessa Su, **Shao-Yuan Lo**, Mingzhe Hu, Yuheng Li, Chih-Wei Chang, Tonghe Wang, Richard L. J. Qiu, and Xiaofeng Yang. “TTT-Vnet: A 3D Vision Test-Time Training Model for Medical Image Analysis.” In *SPIE Medical Imaging (SPIE MI)*, 2025.

- [19] Yuchen Yang, Kwonjoon Lee, Behzad Dariush, Yinzhi Cao, and **Shao-Yuan Lo**. “Follow the Rules: Reasoning for Video Anomaly Detection with Large Language Models.” In *European Conference on Computer Vision (ECCV)*, 2024.
- [18] Hongji Guo, Nakul Agarwal, **Shao-Yuan Lo**, Kwonjoon Lee, and Qiang Ji. “Uncertainty-aware Action Decoupling Transformer for Action Anticipation.” In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. [**Highlight (top 3%)**]
- [17] Himangi Mittal, Nakul Agarwal, **Shao-Yuan Lo**, and Kwonjoon Lee. “Can’t Make an Omelette without Breaking Some Eggs: Plausible Action Anticipation using Large Video-Language Models.” In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.
- [16] Shaoyan Pan, **Shao-Yuan Lo**, Chih-Wei Chang, Ella Salari, Tonghe Wang, Justin Roper, Aparna H. Kesarwala, and Xiaofeng Yang. “Patient-specific 3D Volumetric CBCT Image Reconstruction with Single X-ray Projection Using Denoising Diffusion Probabilistic Model.” In *SPIE Medical Imaging (SPIE MI)*, 2024. [Best Student Paper Award]
- [15] **Shao-Yuan Lo** and Vishal M. Patel. “Adaptive Batch Normalization Networks for Adversarial Robustness.” In *IEEE International Conference on Advanced Video and Signal-based Surveillance (AVSS)*, 2024. [Oral]
- [14] **Shao-Yuan Lo**, Poojan Oza, Sumanth Chennupati, Alejandro Galindo, and Vishal M. Patel. “Spatio-Temporal Pixel-Level Contrastive Learning-based Source-Free Domain Adaptation for Video Semantic Segmentation.” In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.
- [13] Shaoyan Pan, **Shao-Yuan Lo**, Min Huang, Chaoqiong Ma, Jacob Wynne, Tonghe Wang, Tian Liu, and Xiaofeng Yang. “Deep Learning-based Multi-Organ CT Segmentation with Adversarial Data Augmentation.” In *SPIE Medical Imaging (SPIE MI)*, 2023.
- [12] **Shao-Yuan Lo**, Poojan Oza, and Vishal M. Patel. “Adversarially Robust One-class Novelty Detection.” In *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)*, 2022. [**IF=23.6**] [Featured in IEEE CTSoc News on Consumer Technology, April 2024]
- [11] **Shao-Yuan Lo**, Wei Wang, Jim Thomas, Jingjing Zheng, Vishal M. Patel, and Cheng-Hao Kuo. “Learning Feature Decomposition for Domain Adaptive Monocular Depth Estimation.” In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022.
- [10] **Shao-Yuan Lo** and Vishal M. Patel. “Exploring Adversarially Robust Training for Unsupervised Domain Adaptation.” In *Asian Conference on Computer Vision (ACCV)*, 2022.
- [9] **Shao-Yuan Lo** and Vishal M. Patel. “Defending Against Multiple and Unforeseen Adversarial Videos.” In *IEEE Transactions on Image Processing (T-IP)*, 2021. [**IF=11.0**] [Journal presentation at ICIIP 2022]
- [8] **Shao-Yuan Lo** and Vishal M. Patel. “Error Diffusion Halftoning Against Adversarial Examples.” In *IEEE International Conference on Image Processing (ICIP)*, 2021.
- [7] **Shao-Yuan Lo**, Jeya Maria Jose Valanarasu, and Vishal M. Patel. “Overcomplete Representations Against Adversarial Videos.” In *IEEE International Conference on Image Processing (ICIP)*, 2021.
- [6] **Shao-Yuan Lo** and Vishal M. Patel. “MultAV: Multiplicative Adversarial Videos.” In *IEEE International Conference on Advanced Video and Signal-based Surveillance (AVSS)*, 2021.
- [5] **Shao-Yuan Lo**, Hsueh-Ming Hang, Sheng-Wei Chan, and Jing-Jhih Lin. “Efficient Dense Modules of Asymmetric Convolution for Real-Time Semantic Segmentation.” In *ACM International Conference on Multimedia in Asia (MMAsia)*, 2019. [**Best Paper Award**]
- [4] **Shao-Yuan Lo** and Hsueh-Ming Hang. “Exploring Semantic Segmentation on the DCT Representation.” In *ACM International Conference on Multimedia in Asia (MMAsia)*, 2019. [Oral]
- [3] **Shao-Yuan Lo**, Hsueh-Ming Hang, Sheng-Wei Chan, and Jing-Jhih Lin. “Multi-Class Lane Semantic Segmentation using Efficient Convolutional Networks.” In *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2019.
- [2] Shang-Wei Hung, **Shao-Yuan Lo**, and Hsueh-Ming Hang. “Incorporating Luminance, Depth and Color Information by a Fusion-based Network for Semantic Segmentation.” In *IEEE International Conference on Image Processing (ICIP)*, 2019. [Oral]
- [1] Ping-Rong Chen*, **Shao-Yuan Lo***, Hsueh-Ming Hang, Sheng-Wei Chan, and Jing-Jhih Lin. “Efficient Road Lane Marking Detection with Deep Learning.” In *IEEE International Conference on Digital Signal Processing (DSP)*, 2018.

PATENTS

- [P6] Bardia Safaei, **Shao-Yuan Lo**, and Faizan Siddiqui. “Data-efficient Visual Instruction Tuning for Multimodal Large Language Models.” US Patent (filed), 2024.
- [P5] Jiacong Xu, **Shao-Yuan Lo**, and Isht Dwivedi. “Towards Zero-Shot Anomaly Detection and Reasoning with Multimodal Large Language Models.” US Patent (filed), 2024.
- [P4] Chunhui Zhang, **Shao-Yuan Lo**, Kwonjoon Lee, and Nakul Agarwal. “Scaling Situated Reasoning on Multimodal Theory-of-Mind.” US Patent (filed), 2024.
- [P3] Yuxiang Guo, **Shao-Yuan Lo**, Kwonjoon Lee, Faizan Siddiqui, and Enna Sachdeva. “Spatiotemporal Stimuli-aware Video Affective Reasoning with Multimodal Large Language Models.” US Patent (filed), 2024.
- [P2] Yuchen Yang, **Shao-Yuan Lo**, Kwonjoon Lee, and Behzad Dariush. “Follow the Rules: Reasoning for Video Anomaly Detection with Large Language Models.” US Patent (filed), 2024.
- [P1] Himangi Mittal, Nakul Agarwal, **Shao-Yuan Lo**, and Kwonjoon Lee. “Plausible Action Anticipation using Large Video-Language Models.” US Patent (filed), 2024.

INVITED TALKS

- **04/15/2025:** “Anomaly Detection in the Era of Multimodal Large Language Models.” At the University of California, Santa Cruz. Hosts: Prof. Cihang Xie and Prof. Yuyin Zhou.
- **04/12/2025:** “Towards Zero-Shot Anomaly Detection and Reasoning with Multimodal Large Language Models.” At the China Society of Image and Graphics (CSIG) Wuhan Center (Virtual). Host: Prof. Yunkang Cao.
- **03/10/2025:** “Anomaly Detection in the Era of Multimodal Large Language Models.” At the University of California, Riverside. Host: Prof. Amit K. Roy-Chowdhury.
- **03/04/2025:** “Anomaly Detection in the Era of Multimodal Large Language Models.” WACV 2025 Workshop on Challenge Of Out-Of-Label (COOL) in Autonomous Driving. Hosts: Dr. Abhishek Aich and Dr. Ali AlShami.
- **11/04/2024:** “AI Safety and Beyond: Robustness, Monitoring, and Alignment.” At National Taiwan University, Taiwan. Host: Prof. I-Hsiang Wang.
- **09/19/2024:** “Multimodal Large Language Models for Driving Safety Applications.” At Jilin University, China (Remote). Host: Prof. Hongxia Xie.
- **01/09/2024:** “Robust Computer Vision Against Adversarial Examples and Domain Shifts.” At National Yang Ming Chiao Tung University, Taiwan. Host: Prof. Wen-Hsiao Peng.
- **12/14/2023:** “Robust Computer Vision Against Adversarial Examples and Domain Shifts.” At Academia Sinica, Taiwan. Host: Prof. Jun-Cheng Chen.
- **04/14/2023:** “Robust Computer Vision Against Adversarial Examples and Domain Shifts.” At Computational Cognition, Vision, and Learning (CCVL) Lab, Johns Hopkins University. Host: Prof. Alan Yuille.
- **01/10/2022:** “Defending Against Multiple and Unforeseen Adversarial Videos.” At National Yang Ming Chiao Tung University, Taiwan. Host: Prof. Wen-Hsiao Peng.
- **01/05/2022:** “Defending Against Multiple and Unforeseen Adversarial Videos.” At Academia Sinica, Taiwan. Host: Prof. Jun-Cheng Chen.
- **06/19/2021:** “Adversarial Attacks and Defenses in Videos.” At CVPR 2021 Tutorial on Adversarial Machine Learning in Computer Vision (Virtual). Host: Prof. Cihang Xie.

ACADEMIC ACTIVITIES

- **Tutorial Chair:** AVSS 2025
- **Associate Editor:** ICRA 2025
- **Area Chair:** CVPR 2024 AI4CG Workshop
- **Journal Reviewer:** IEEE T-PAMI, IEEE T-IP, IEEE RA-L, IEEE T-CSVT, IEEE T-SMC, IEEE JETCAS, Pattern Recognition, Neurocomputing, Medical Physics

- **Conference Reviewer:** CVPR (2022-25), ICCV (2021-25), ECCV (2022-24), NeurIPS (2025), ICLR (2023), ICML (2025), AAAI (2023-24), IJCAI (2025), ACM MM (2025), WACV (2021-25), ACCV (2022-24), ICIP (2022-24), ICPR (2024), AVSS (2021-22)
- **Teaching Assistant:** Deep Learning (EN.520.638), JHU, Spring (2021, 2022, 2023)

MENTORING

Research Interns at Honda Research Institute USA

- Chunhui Zhang (PhD student, Dartmouth), published a paper at ICML 2025 [26] Summer 2024
- Jiacong Xu (PhD student, JHU), published a paper at CVPR 2025 [24] Summer 2024
- Bardia Safaei (PhD student, JHU), published a paper at CVPR 2025 [25] Summer 2024
- Yuxiang Guo (PhD student, JHU), submitted a paper to IJCV [28] Spring 2024
- Yuchen Yang (PhD student, JHU), published a paper at ECCV 2024 [19] Fall 2023

Research Collaboration Mentees

- Udit Ghosh (PhD student, UC Riverside), working on a paper Spring 2025
- Kuang-Ming Chen (MS student, University of Washington), submitted a paper to ICCV [27] Spring 2025
- Jr-Jen Chen (MS student, NTU), submitted a paper to ICCV [27] Spring 2025
- Zhuoli Zhuang (PhD student, University of Technology Sydney), working on a paper 2024-2025
- Yu-Jen Tseng (MS student, NYCU), working on a paper 2024-2025
- Anirudh Thatipelli (MS student, UC Riverside), published a paper at CVIU [22] Spring 2024
- Cheng Chien (MS student, NYCU), published a paper at ICLR 2025 [21] Spring 2024
- Shaoyan Pan (PhD student, Emory), published four papers at T-MI [21] and SPIE MI 2023-25 [13, 16, 20] 2022-2024
- Shang-Wei Hung (BS student, NCTU), published a paper at ICIP [2] Summer 2018

PROGRAMMING SKILLS

- **Programming Languages:** Python, MATLAB, C/C++
- **Deep Learning Libraries:** PyTorch, TensorFlow, Caffe
- **Hardware Design Tools:** Verilog, HSPICE, Cadence Virtuoso

LEADERSHIP

- Vice President, JHU Taiwanese Student Association 2020 –2022
- Secretary, NCTU EECS Student Association 2015 –2016
- Treasurer, NCTU Chinese Chess Club 2014 –2015
- Arts Chair, NCTU EECS Summer Camp for High School Students 2014 –2015

PRESS COVERAGE

- **10/11/2022:** I was reported on Taiwan's *Formosa Television*, *EBC News* and *New Tang Dynasty Television* while serving as Vice President of the JHU Taiwanese Student Association.
- **06/01/2019:** I was reported on Taiwan's *Liberty Times*, *United Daily News* and *Commercial Times* when I graduated from NCTU.
- **07/19/2013:** I was reported on Taiwan's *China Times* when I graduated from senior high school (National Experimental High School at Hsinchu Science Park).
- **06/04/2010:** I was reported on Taiwan's *Liberty Times* when I graduated from junior high school.