

Syed Hesham

PhD Candidate, Electrical & Electronic Engineering, NTU Singapore | A*STAR CIS Scholar

✉ syedhesh002@e.ntu.edu.sg 🔗 linkedin.com/in/syedhesham 📄 github.com/Ashesham
🌐 ashesham.github.io 📄 Google Scholar

PROFESSIONAL SUMMARY

As a PhD candidate and A*STAR CIS Scholar at NTU Singapore, I study multimodal video understanding, curious what a model can really afford to remember. That curiosity shapes how I build, using state-space and spatiotemporal-compression to let vision-language foundation models perceive and reason over long videos, published first-author at CVPR 2025 and ECCV 2026. I believe ideas sharpen through two-way exchange, so my strongest work happens thinking alongside others.

EDUCATION

Nanyang Technological University (NTU), Singapore <i>PhD in Electrical & Electronic Engineering</i> A*STAR CIS Scholar Expected Dec 2027 GPA: 5.0/5.0. Topic: Effective Visual Perception with Large Foundation Models.	Jan 2024 – Present
Nanyang Technological University (NTU), Singapore <i>BEng in Electrical & Electronic Engineering</i> GPA: >4.9/5, Dean's List (x4) Final year project: Real-time SLAM using deep features.	Aug 2020 – Dec 2023
Republic Polytechnic, Singapore <i>Diploma in Electrical & Electronic Engineering</i> Director's Roll of Honour	2016 – 2019

WORK EXPERIENCE & RESEARCH PROJECTS

Nanyang Technological University / A*STAR I2R, Singapore <i>Doctoral Researcher</i> - Research focus: Efficient Video Semantic Segmentation, Video Reasoning Segmentation, and Multimodal Video Understanding with large foundation models (CVPR 2025, ECCV 2026, MIR). - Faced with segmentation models that are too costly for long video, designed TV3S (Temporal State Space Sharing) for video semantic segmentation (CVPR 2025) and STAC (Selective Spatiotemporal Aggregation and Compression) for multimodal LLM-based video reasoning segmentation (ECCV 2026). - Benchmarked SAM2 video foundation models for video semantic segmentation (MIR) and co-authored a survey on video scene parsing (under review).	Jan 2024 – Present
Nankai University, China <i>Visiting Researcher – Prof. Yun Liu's group, enhancing multimodal understanding in video models</i>	Dec 2025 – Feb 2026
Republic Polytechnic (SEG), Singapore <i>Research Assistant</i> - Developed a cost-effective near-field pose estimation system that secured over S\$200K in government funding, and a real-time SLAM visualization pipeline in Python, C++, C#, and ROS. - Mentored 40+ students across project teams, leading them to national competition wins (Engineering Innovation Challenge 2022, IDM Smart Nation Award 2021) and to peer-reviewed publications.	Nov 2020 – Jan 2024
Continental Automotive Pte. Ltd. (R&D), Singapore <i>Research Intern</i> Delivered practical solutions across three internal R&D projects and presented results to cross-functional teams.	Mar 2018 – Aug 2018

SELECTED PUBLICATIONS

- Exploiting Temporal State Space Sharing for Video Semantic Segmentation.** CVPR 2025. First author.
- STAC: Selective Spatiotemporal Aggregation and Compression for Video Reasoning Segmentation.** ECCV 2026. First author.
- Evaluating SAM2 for Video Semantic Segmentation.** Machine Intelligence Research (MIR), accepted. First author.

SKILLS

Technical: Python, PyTorch, C/C++, TensorFlow, Distributed Training (Multi-node, Multi-GPU), Git, Docker, Linux, ROS
Research: Multimodal Video Understanding, Video Reasoning & Segmentation, Multimodal LLMs & VLMs (incl. task-specific fine-tuning), Vision-Language Alignment, State Space Models (Mamba), Spatiotemporal Compression, Model Efficiency and Optimization, SLAM & 3D Perception
Transferable: Technical Writing and Publication, Mentorship and Team Leadership, Cross-functional Collaboration

ADDITIONAL INFORMATION

Awards: Lambda Research Grant (2025); IEEE ICME Grand Challenge, Bronze, Team Leader (2021); IES National Innovation Challenge, Bronze (2019)
Languages: English (fluent), Tamil (fluent), Hindi (fluent)
Leadership: NTU Graduate Students' Association – Assistant President (2025), Assistant Vice-President Academic (2024)

Syed Hesham · Publications

PUBLICATIONS

[1] **Exploiting Temporal State Space Sharing for Video Semantic Segmentation**

Syed Ariff Syed Hesham, Yun Liu, Guolei Sun, Henghui Ding, Jing Yang, Ender Konukoglu, Xue Geng, Xudong Jiang
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025. Acceptance rate 22.1%.

[2] **STAC: Selective Spatiotemporal Aggregation and Compression for Video Reasoning Segmentation**

Syed Ariff Syed Hesham, Yun Liu, Guolei Sun, Jing Yang, Henghui Ding, Xue Geng, Xudong Jiang
European Conference on Computer Vision (ECCV), 2026.

[3] **Evaluating SAM2 for Video Semantic Segmentation**

Syed Hesham Syed Ariff, Yun Liu, Guolei Sun, Jing Yang, Henghui Ding, Xue Geng, Xudong Jiang
Machine Intelligence Research (MIR), accepted 2026. [arXiv:2512.01774](#).

[4] **A Comprehensive Survey on Video Scene Parsing: Advances, Challenges, and Prospects**

Guohuan Xie, **Syed Ariff Syed Hesham**, Wenya Guo, Bing Li, Ming-Ming Cheng, Guolei Sun, Yun Liu
arXiv preprint arXiv:2506.13552, 2025. Under review.

[5] **Tracking and Monitoring of Underwater Object with SLAM**

Lijun Jiang, **Syed Ariff Syed Hesham***, Lan JunHang, Seah Kai Wen Kelvin, Yuhan Jiang, Yao Mengdi, Wei Dongliang, Bo Jiang
IEEE 19th Conference on Industrial Electronics and Applications (ICIEA), 2024.

[6] **Adapted Lightweight MobileNet for Tire Pattern Classification**

Syed Ariff Syed Hesham, Lijun Jiang, Keng Pang Lim, Sui JinZhou, Phua Zen Long, Lan JunHang, Zhao Hu, Yang Shanglin, Zhao Xu, Zhao Chuanfeng
IEEE 18th Conference on Industrial Electronics and Applications (ICIEA), 2023.

[7] **Leveraging on Few-Shot Learning for Tire Pattern Classification in Forensics**

Lijun Jiang, **Syed Ariff Syed Hesham***, Keng Pang Lim, Changyun Wen
Journal of Automation and Intelligence, 2023.

[8] **A Versatile Application for Visual Simultaneous Localization and Mapping with Object Detection**

Lijun Jiang, **Syed Ariff Syed Hesham***, Keng Pang Lim, Yusong Wang, Hongkai Lin, Yuhang Zhao
IEEE 17th Conference on Industrial Electronics and Applications (ICIEA), 2022.

[9] **YOLO Based Thermal Screening Using Artificial Intelligence (AI) for Instinctive Human Facial Detection**

Lijun Jiang, **Syed Ariff Syed Hesham***, Keng Pang Lim, Krishnadas Manoj, Mohammed Razi, Zhang Zishuo, Ma Chenxin, Bo Jiang, Hamid Saeedipour
IEEE 17th Conference on Industrial Electronics and Applications (ICIEA), 2022.

[10] **Borescope Tracking and Visualization of Internal Aero-Structure with VSLAM**

Lijun Jiang, **Syed Ariff Syed Hesham***, Keng Pang Lim, Sui JinZhou, Bo Jiang
IEEE 17th Conference on Industrial Electronics and Applications (ICIEA), 2022.

[11] **Improving Recognition Performance for Low-Resolution Images Using DBPN**

Lijun Jiang, Keng Pang Lim, **Syed Ariff Syed Hesham***
IEEE 16th Conference on Industrial Electronics and Applications (ICIEA), 2021.

[12] **Crack Detection on Aircraft Composite Structures Using Faster R-CNN**

Lijun Jiang, **Syed Ariff Syed Hesham***, Haoming Shi, Hamid Saeedipour
IEEE 16th Conference on Industrial Electronics and Applications (ICIEA), 2021.

[13] **PIX2PT Map for Transfer-Based Few-Shot Learning**

Syed Ariff Syed Hesham, Sui JinZhou*, Gabrielle Ee Song Xin, Lek Chen Ping, Lijun Jiang
IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2021.

* Equal contribution.