

To Thanh Dat

✉ ttdat2419@clc.fitus.edu.vn | 🌐 github.com/tothanhdatt2006

Education

University of Science - VNUHCM

2024 - Present

- Bachelor in Information Technology. GPA 3.75/4.0

Projects

Traffic Video Captioning & VQA | 🌐

May 2025 - Present

- Fine-tuned **Qwen2.5-VL-7B** using **LoRA** and **DoRA** to enhance temporal and spatial understanding for traffic video captioning and visual question answering (VQA)
- Designed domain-specific textual prompts that improved average validation scores by **+9.294** compared to baseline inference without prompts
- Ranked **7th** in the **9th AI City Challenge** at **ICCV 2025** with a benchmark score of **55.655**

Blast Fragment Segmentation | 🌐

March 2025 - May 2025

- Fine-tuned a **Mask R-CNN** model on a dense post-blast rock fragmentation dataset to perform instance segmentation
- Achieved **0.9080 mIoU** on the validation set and **0.8865 mIoU** on the testing set
- Designed a post-processing algorithm to estimate rock fragment sizes from predicted segmentation masks, enabling automated fragmentation analysis

Publications

(*Co-first author)

- Tinh-Anh Nguyen-Nhu*, Triet Dao Hoang Minh*, **Dat To-Thanh***, Phuc Le-Gia, Tuan Vo-Lan and Tien-Huy Nguyen. "STER-VLM: Spatio-Temporal With Enhanced Reference Vision-Language Models". In IEEE International Conference on Computer Vision (ICCV), 2025. [arXiv:2508.13470](https://arxiv.org/abs/2508.13470)

Activities

Google Developer Club on Campus - HCMUS

September 2024 - Present

- Contributed ideas for large workshops.
- Composed drafts and collaborated in organizing a seminar.

Skills & Interests

Languages Python, C/C++

Libraries Pytorch, Torchvision

Languages English, Vietnamese