Phuc, Nguyen Duc Anh

phucnda.github.io | @ phucnda (at) gmail (dot) com | phucnda

Education

University of Maryland, College Park — Maryland, USA
Ph.D. in Computer Science – supervised by Prof. Ming C. Lin — *Sep 2025 - Now*

University of Information Technology - Vietnam National University — Ho Chi Minh, Vietnam
B.Sc. in Computer Science — *Nov 2020 – Aug 2023*

High School for the Gifted - Vietnam National University — Ho Chi Minh, Vietnam
Specialized in Physics — *Sep 2017 – July 2020*

Research Interests

My research focuses on Computer Vision, particularly on leveraging Vision-Language Models to achieve a comprehensive understanding of 2D and 3D environments without vocabulary constraints. I am also dedicated to integrating Multimodal Large Language Models to enable more interactive and intuitive communication between humans and machines. My work involves developing algorithms that extract both geometric and semantic information from real-world settings, with applications in robotics, autonomous driving, and augmented reality.

Experience

University of Maryland, College Park – GAMMA Lab — **Maryland, USA**
PhD Student – Supervised by Prof. Ming C. Lin — *Sep 2025 – Now*

- Research Project: Open-World Visual Odometry
  - Developing a temoporal and spatial aware visual odometry system estimating real-world–scale ego-motion from any monocular dashcam footage captured by any camera. One paper is currently under review.
- Research Project: High-Quality 3D Reconstruction
  - Piror works estimate only local perspectives from local observations. Can a reliable ego-trajectory serve as a backbone to fuse these local views into a globally consistent world model? Given the strong priors encoded in modern foundation models, do we still need end-to-end retraining for scene reconstruction? Project is ongoing.

SpreeAI — **California, USA**
AI Research Engineer – In close collaboration with Dr. Aayush Bansal and Dr. Minh Vo — *Nov 2024 – Sep 2025*

- Research Project: Multimodal representations for physically-grounded Multimodal-LLM.
  - How can we determine whether a multimodal LLM's grounding is truly reliable, given that even humans often do not know what they do not know? Can a model assess whether it understands or is hallucinating? We address this through an Active Continual Learning framework with human-in-the-loop feedback, where the model actively selects the most informative data and learns incrementally rather than retraining from scratch. One internal system delivered.

VinAI Research – Qualcomm — **Hanoi, Vietnam**
AI Research Resident – Supervised by Dr. Anh Tran and Prof. Cuong Pham — *Feb 2023 – Nov 2024*

- Research Project: Open-Vocabulary 3D Instance Segmentation
  - Developing an algorithm for tackling the open-vocabulary 3D point cloud instance segmentation by using 2D prior, achieves state-of-the-art results on five different dataset benchmarks. One paper accepted to the CVPR'24, winning two workshop challenges ICCV'23, and CVPR'24.
- Research Project: Class-agnostic 3D Instance Segmentation
  - Developing an algorithm for high-quality class-agnostic 3D instance segmentation by leveraging a tracking foundation model and an optimization-based mask aggregation approach. One paper accepted to the CVPR'25.
- Research Project: Vocabulary-Free 3D Point Cloud Understanding
  - Introducing a new benchmark and algorithms for open-ended 3D point cloud instance segmentation, leveraging Large-Language Models to establish several baselines. One paper accepted to the ICCV'25.
- Applied R. Project: Vietnam Geographic Map Understanding – *Advised by Prof. Minh Hoai*

* VinAI Research collaborates with the Ministry of Information and Communications on a project utilizing computer vision to identify counterfeit Vietnam geographic maps online, facilitated by the introduction of the extensive VinMap dataset. One paper accepted to the Vietnam local conference.

University of Information Technology – Vietnam National University **Ho Chi Minh, Vietnam**
Research BS Student *Nov 2020 – Aug 2023*

- Graduate Thesis: Hybrid-Anchor Rotation Detector for Oriented Object Detection (9.8/10.0)
  * Developing a hybrid model for Oriented Object Detection achieves state-of-the-art performance with the least training resource on three dataset benchmarks. Best thesis award - one paper accepted to ICCV'25.
- Research Project: Aerial Oriented Object Detection
  * Investigating Oriented Object Detection in Aerial Images. Algorithms implementations and dataset development for adverse weather conditions. Four papers accepted at the local conferences/journals.
- Research Project: Page Object Detection
  * Evaluating the new YOLO design on two dataset benchmarks of page object detection. One paper accepted at the local conferences.

Selected Publications. See full list at Google Scholar

- **Phuc Nguyen***, Anh Nhu*, Ming C. Lin, "OpenVO: Open-World Visual Odometry with Temporal Dynamics Awareness", in *Under Submission.*
- **Phuc Nguyen***, Minh Luu*, Anh Tran, Cuong Pham, Khoi Nguyen, "Open-Ended 3D Point Cloud Instance Segmentation", in *International Conference on Computer Vision (**ICCV**) Workshops, 2025.*
- **Phuc Nguyen**,"HA-RDet: Hybrid Anchor Rotation Detector for Oriented Object Detection", ***Bachelor's Thesis, 2022 - Best Thesis Award.***
  *International Conference on Computer Vision (**ICCV**) Workshops, 2025.*
- **Phuc Nguyen**, Minh Luu, Anh Tran, Cuong Pham, Khoi Nguyen, "Any3DIS: Class-Agnostic 3D Instance Segmentation by 2D Mask Tracking", in *Computer Vision and Pattern Recognition Conference (**CVPR**), 2025.*
- **Phuc Nguyen***, Tuan Duc Ngo*, Evangelos Kalogerakis, Chuang Gan, Anh Tran, Cuong Pham, Khoi Nguyen, "Open3DIS: Open-Vocabulary 3D Instance Segmentation with 2D Mask Guidance", in *Computer Vision and Pattern Recognition Conference (**CVPR**), 2024.*
- ..., **Phuc Nguyen**, Khoi Nguyen, Anh Tran, Cuong Pham,..., "OpenSUN3D: 1st Workshop Challenge on Open-Vocabulary 3D Scene Understanding", in *International Conference on Computer Vision (**ICCV**) Workshops, 2023.*

Awards & Achievements

UMD Dean' Fellowship: Awarded to candidates with exceptional academic records. (2025–2027)
First Prize CVPR Workshop: VinAI-3DIS ranked top-1 in OpenSUN3D CVPR workshop. (2024)
Second Prize ICCV Workshop: VinAI-3DIS ranked top-2 in OpenSUN3D ICCV workshop. (2023)
Best Thesis Award: Awarded to thesis with the highest grade. (2023)
Third Prize UIT AI Challenge: The team ranked top-3 in Scene Text recognition challenge. (2023)
Second Prize UCPC: Ranked top-2 in UIT Collegiate Programming Contest. (2022)
Expert Codeforces: Reaching Expert title on Codeforces – Competitive Programming platform. (2022)
First Prize UIT-AlgoBootcamp: Winning Competitive Programming Competition at UIT. (2021)
Outstanding Student Scholarship: Awarded to students with the best academic performance. (2021)
Outstanding Student in Physics: Awarded to students with the highest GPA in Physics. (2020)

Service

Reviewer:

* IEEE/CVF Computer Vision and Pattern Recognition (CVPR'24, CVPR'26)
* IEEE/CVF International Conference on Computer Vision (ICCV'25)
* European Conference on Computer Vision (ECCV'24)
* International Conference on Learning Representations (ICLR'25)
* Neural Information Processing Systems (NeurIPS'24)
* British Machine Vision Conference (BMVC'25)
* Transactions on Machine Learning Research (TMLR)

Teaching Assistant UMD:

* CMSC420 (Fall'25): Advanced Data Structures