

TRAN DINH PHU

Google Scholar ◇ Homepage ◇ Github ◇ LinkedIn
Phone: (+84) 32-866-6370 ◇ Email: phutx2000@kaist.ac.kr

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST) *May 2028 (Expected)*

Ph.D. in Computer Science
GPA: 3.90/4.3 (Current GPA)
Advisor: Prof. Dr. Daeyoung Kim

Korea Advanced Institute of Science and Technology (KAIST) *February 2025*

M.S. in Computer Science
GPA: 3.94/4.3
Thesis: 윈도우 어텐션을 활용한 초고해상도 이미지 변환 딥러닝 기법
Advisor: Prof. Dr. Daeyoung Kim

Hanoi University of Science and Technology (HUST) *June 2022*

B.E. in Control Engineering and Automation
GPA: 3.66/4.0 (*Top 1%*)
Thesis: Trans2Unet: Neural fusion for nuclei semantic segmentation (9.5/10 - Best Presentation Award)
Advisors: Assoc. Prof. Dr. Thi-Thao Tran and Assoc. Prof. Dr. Van-Truong Pham

Relevant Coursework: Computer Vision | Deep Learning | Advanced Deep Learning | Artificial Intelligence and Machine Learning | Audio-Visual Language Multi-Modal AI | Computational Linguistics

EXPERIENCE & RESEARCH INTERESTS

Image and Video Restoration/Enhancement

Image and Video Super-Resolution/Restoration, Image Enhancement, and Low-Level Vision. Research interests include efficient restoration architectures, diffusion-based generative SR focusing on *addressing the trade-off between perceptual and fidelity quality*, temporal consistency modeling, and foundation models for image and video enhancement.

Vision-Audio Multimodal Learning

Audio-Visual Question Answering, Long-Form Video Understanding, Multimodal Reasoning, and Agentic AI Systems. Research focuses on reliable multimodal learning, parameter-efficient adaptation of foundation models, cross-modal representation learning, and scalable reasoning over long-horizon audio-video content.

Training Method for Model Generalization, Regularization

Model generalization and robust optimization. Super interested in *Sharpness-Aware Minimization (SAM)*, federated learning, and regularization techniques such as SWA for improving robustness, fairness, and training stability under limited data and compute budgets.

PUBLICATIONS

Knowing When to Answer: Adaptive Confidence Refinement for Reliable Audio-Visual Question Answering

Dinh Phu Tran, Jihoon Jeong, Saad Wazir, Seongah Kim, Thao Do, Cem Subakan, Daeyoung Kim
Under review

TB-AVA: Text as a Semantic Bridge for Audio-Visual Parameter Efficient Finetuning

Seongah Kim, Dinh Phu Tran, Hyeontaek Hwang, Saad Wazir, Duc Do Minh, Daeyoung Kim
Under review

FreqOrtho-SR: Frequency-Guided Orthogonal Expert Learning for Real-World Image Super-Resolution

Minh Son Hoang*, **Dinh Phu Tran***, Quyen Nguyen Duc, Dam Hoang Phuong, Daeyoung Kim
ECCV 2026, Malmö, Sweden (Rank A)*

MedCAGD: Context-Aware Gated Decoder for Robust Medical Image Segmentation
Saad Wazir, Patrick Vibild, **Dinh Phu Tran**, Seongah Kim, Daeyoung Kim
ECCV 2026, Malmö, Sweden (Rank A)*

Efficient Encoder-Only Context Compression via Marginal Contribution Scoring
Thao Do, **Dinh Phu Tran**, An Vo, Seon Kwon Kim, Daeyoung Kim
ICML 2026 Workshop, Seoul, South Korea (-)

SAT: Selective Aggregation Transformer for Image Super-Resolution
Dinh Phu Tran, Thao Do, Saad Wazir, Seongah Kim, Seon Kwon Kim, Daeyoung Kim
CVPR 2026 (Finding Track), Denver, USA (Rank A)*

VSRM: A Robust Mamba-Based Framework for Video Super-Resolution
Dinh Phu Tran, Dao Duy Hung, Daeyoung Kim
ICCV 2025, Hawaii, USA (Rank A)*

Reference-Based Post-OCR Processing with LLM for Precise Diacritic Text in Historical Document Recognition
Thao Do, **Dinh Phu Tran**, An Vo, Daeyoung Kim
AAAI 2025 (Social Track), Philadelphia, Pennsylvania, USA (Rank A)*

Channel-Partitioned Windowed Attention And Frequency Learning for Single Image Super-Resolution
Dinh Phu Tran, Dao Duy Hung, Daeyoung Kim
BMVC 2024, Glasgow, UK (Rank A)

PATENT

Video Super-Resolution Technique Leveraging Spatial-Temporal Mamba Architecture
Under review

HONORS AND AWARDS

Huynhai Global Scholarship for Graduate Student (5 slots/admission cycle)	2025-2029
Kaist Scholarship for PhD Program (Full tutions and fees)	2025-2029
Kaist Scholarship for Master Program (Full tutions and fees)	2023-2024
Best Bachelor Thesis Award	09/2022
Scholarship Encouraging Study (Level A - Highest Level)	2021-2022
Scholarship Encouraging Study (Level A - Highest Level)	2020-2021
Nitori International Scholarship	2019-2020
Graduated as valedictorian from high school	09/2018

SKILLS

Programming Language	Python, Bash, C/C++
Library	PyTorch, HuggingFace, Tensorflow, Pandas, Numpy, OpenCV
Deep Learning	Neural Nets (CNN, Transformer, Mamba,...), Generative Models,...
Version Control Systems	Git
Language	English, Vietnamese (mother tongue)

ACADAMIC SERVICES

Reviewer for AAAI 2026, CVPR 2026, ICML 2026, NeurIPS 2026