

Thanh Van Trong Tran

thanhtvt.github.io | trantrongthanhhp@gmail.com | linkedin.com/in/thanhtvt | github.com/thanhtvt

EDUCATION

Nanyang Technological University (NTU)

Singapore

Ph.D. in Electrical and Electronic Engineering

Aug 2026 – Present

- Advisor: [Prof. Woon-Seng Gan](#)

VNU University of Engineering and Technology (VNU-UET)

Hanoi, Vietnam

Bachelor in Computer Science (Honors)

Aug 2019 – Jun 2023

- Thesis: Automatic Speech Recognition in Low-Resource Condition
- GPA: 3.71/4.00 (top 5%)

EXPERIENCE

AI Research Resident

Aug 2023 – Jul 2026

FPT Software – AI Center

Hanoi, Vietnam

- Researched speech AI and multimodal learning for emotion recognition, vision-guided audio and speech synthesis.
- Developed black-box optimization methods in discrete and latent spaces to improve protein sequences.
- Supervisor: [Prof. Truong-Son Hy](#) and [Dr. Van Nguyen](#).

Student Researcher

Sep 2022 – May 2023

Institute for Artificial Intelligence (IAI), VNU-UET

Hanoi, Vietnam

- Researched Person/Vehicle Re-identification (ReID) and Instance Retrieval problems.
- Developed an ESPNet-like speech processing toolkit in TensorFlow 2, streamlining development and deployment.
- Supervisor: [Dr. Long Quoc Tran](#)

Speech Processing Intern

Oct 2021 – Nov 2022

VinBigData – Spoken Language Processing Department

Hanoi, Vietnam

- Designed a C-based telephony speech augmentation pipeline, reducing WER by 25% on production benchmarks.
- Developed a neural decoder that improved decoding accuracy by 10% over vanilla beam search.
- Delivered a low-latency streaming ASR model (~70ms) with approx. 10% WER for in-car assistant.
- Optimized the in-house ASR framework, improving scalability and processing performance.
- Developed a specialized QA app for the VinFast's Virtual Assistant, accelerating the EV deployment pipeline.

SELECTED PUBLICATIONS ([GOOGLE SCHOLAR](#))

DiFlowDubber: Discrete Flow Matching for Automated Video Dubbing via Cross-Modal Alignment and Synchronization

Ngoc-Son Nguyen, [Thanh V. T. Tran](#), Jeongsoo Choi, Hieu-Nghia Huynh-Nguyen, Truong-Son Hy, Van Nguyen

Computer Vision and Pattern Recognition Conference (CVPR), 2026. [\[paper\]](#) [\[code\]](#) [\[website\]](#)

Speech Reconstruction from Silent Videos via Acoustic-Semantic Decomposed Modeling

Long-Khanh Pham, [Thanh V. T. Tran](#), Minh-Tan Pham, Van Nguyen

Interspeech, 2025. [\[paper\]](#) [\[code\]](#) [\[website\]](#)

Effective Context Modeling Framework for Emotion Recognition in Conversations

Cuong Tran Van, [Thanh V. T. Tran*](#), Van Nguyen, Truong-Son Hy*

International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2025. [\[paper\]](#) [\[code\]](#) [\[poster\]](#)

GROOT: Effective Design of Biological Sequences with Limited Experimental Data

[Thanh V. T. Tran](#), [Nhat Khang Ngo*](#), Viet Anh Nguyen, Truong-Son Hy*

Conference on Knowledge Discovery and Data Mining (KDD), 2025. [\[paper\]](#) [\[code\]](#) [\[poster\]](#) [\[slides\]](#)

PREPRINTS

Precise Video-to-Audio Generation with Cross-Modal Alignment in Latent Space

Thanh V. T. Tran, Ngoc-Son Nguyen, Luong Tran, Long-Khanh Pham, Paarth Neekhara, Shehzeen Hussain, Van Nguyen

DiFlow-TTS: Compact and Low-Latency Zero-Shot Text-to-Speech with Factorized Discrete Flow Matching

Ngoc-Son Nguyen, Thanh V. T. Tran, Hieu-Nghia Huynh-Nguyen, Truong-Son Hy, Van Nguyen
arXiv, 2025. [\[paper\]](#) [\[code\]](#) [\[website\]](#)

MASTRO: Hierarchical Flow Matching for Semantic-to-Acoustic Speech Token Reconstruction from Video

Long-Khanh Pham, Ngoc-Son Nguyen, Jeongsoo Choi, Hieu-Nghia Huynh-Nguyen, Thanh V. T. Tran, Minh-Tan Pham, Van Nguyen

HONORS AND AWARDS

The Rising Innovator

FPT Software – AI Center

Apr 2026

- Achieved *Third Place* among 20 finalists in the internal AI Innovation Hackathon Q1 2026.

Best Performance Award 2025

FPT Software – AI Center

Jan 2026

- Recognized as one of the top 3 performers among 100+ AI engineers and researchers in 2025.

Second Prize

Association for Vietnamese Language and Speech Processing

2022

- Achieved *Second Place* in Task 1: Automatic Speech Recognition at VLSP 2022, surpassing 30+ teams nationwide

Third Prize

Association for Vietnamese Language and Speech Processing

2022

- Achieved *Third Place* in Task 2: Automatic Speech Recognition at VLSP 2022, surpassing 30+ teams nationwide.

SKILLS

Languages: Vietnamese (Native), English (IELTS 7.5)

Programming and Software: Python, C/C++, Bash, Git, L^AT_EX, Docker, Linux

Machine Learning Package: PyTorch, TensorFlow, Transformers, Scikit-Learn, NVIDIA-NeMo

ACTIVITIES

Data Science Mentor

Math and Summer Science Program (MaSSP)

2023

- Lecturer at MaSSP (a non-profit summer STEM program for Vietnamese high school students), teaching Calculus and Foundations of Computer Vision.