THANH VAN TRONG TRAN

(+84) 989955612 ⋄ Hanoi, Vietnam

[Homepage](#) ⋄ [Email](#) ⋄ [LinkedIn](#) ⋄ [GitHub](#) ⋄ [Google Scholar](#)

RESEARCH & WORK EXPERIENCE

AI Research Resident — Aug 2023 - Present
FPT Software — *largest IT services company in Southeast Asia* — *Hanoi, Vietnam*

- Developed an end-to-end speech-to-text summarization pipeline for doctor-patient dialogues, enabling automatic generation of SOAP notes.
- Explored speech AI and multimodal learning for diverse applications, including multisensory emotion recognition, vision-guided speech and audio synthesis, text-to-speech, and LLM-agent development.
- Developed methods to improve protein characteristics at the sequence level. My research investigated black-box optimization in latent space to effectively guide the optimization trajectory of protein sequences.
- Advisors: **Assistant Prof. Truong-Son Hy** and **Dr. Van Nguyen**.

Student Researcher — Sep 2022 - May 2023
Institute for Artificial Intelligence — *VNU University of Engineering and Technology, Hanoi*

- Built and made open-source the error analysis tool for multi-target multi-camera tracking system (*[GitHub](#)*).
- Developed graph neural network-based solution for re-identification problem in multi-target multi-camera tracking system.
- Advisor: **Dr. Long Quoc Tran**.

Speech Processing Intern — Oct 2021 - Nov 2022
VinBigData — *Conversational AI solutions* (now VinSmart Future) — *Hanoi, Vietnam*

- Designed a telephony-speech augmentation pipeline in C, achieving a 25% reduction in WER on production benchmarks.
- Developed a neural decoder that outperforms vanilla beam search by 10%.
- Delivered a low-latency streaming ASR ($\approx$ 70ms end-to-end, 12% WER on in-house dataset) for automotive deployment.
- Implemented and maintained company's in-house speech recognition framework.

EDUCATION

Bachelor of Computer Science, VNU University of Engineering and Technology — Aug 2019 - Jun 2023
Top 10% of the entire graduating cohort — GPA: 3.71/4.0

PUBLICATIONS

(*) Equal contribution

Conference Papers

Speech Reconstruction from Silent Videos via Acoustic-Semantic Decomposed Modeling | [Link](#)
Long-Khanh Pham, **Thanh V. T. Tran**, Minh-Tan Pham, Van Nguyen.
Interspeech 2025.

Effective Context Modeling Framework for Emotion Recognition in Conversations | [Link](#)
Cuong Tran Van*, **Thanh V. T. Tran***, Van Nguyen, Truong-Son Hy.
ICASSP 2025.

GROOT: Effective Design of Biological Sequences with Limited Experimental Data | [Link](#)
Thanh V. T. Tran*, Nhat Khang Ngo*, Viet Anh Nguyen, Truong-Son Hy.
KDD 2025.

Journal Papers

Latent-based Directed Evolution accelerated by Gradient Ascent for Protein Sequence Design | [Link](#)
Thanh V. T. Tran, Nhat Khang Ngo, Viet Thanh Duy Nguyen, Truong-Son Hy.
Machine Learning: Science and Technology (Q1, IF: 4.6).

Protein Design by Directed Evolution guided by Large Language Models | [Link](#)
Thanh V. T. Tran, Truong-Son Hy.
IEEE Transactions on Evolutionary Computation (Q1, IF: 12).

Precise Video-to-Audio Generation with Cross-Modal Alignment in Latent Space
<u>Thanh V. T. Tran</u>, Ngoc-Son Nguyen, Luong Tran, Long-Khanh Pham, Paarth Neekhara, Shehzeen Hussain, Van Nguyen.

DiFlowDubber: Discrete Flow Matching for Automated Video Dubbing via Cross-Modal Alignment and Synchronization
Ngoc-Son Nguyen, <u>Thanh V. T. Tran</u>, Jeongsoo Choi, Hieu-Nghia Huynh-Nguyen, Truong-Son Hy, Van Nguyen.

DiFlow-TTS: Compact and Low-Latency Zero-Shot Text-to-Speech with Factorized Discrete Flow Matching
Ngoc-Son Nguyen, <u>Thanh V. T. Tran</u>, Hieu-Nghia Huynh-Nguyen, Truong-Son Hy, Van Nguyen.

AgentSound: A LLM-Driven Multi-Agent Framework for Augmenting Sounds with Semantic Context Guidance
Van Truong Hoang, Viet Pham, <u>Thanh V. T. Tran</u>, Van Nguyen, Harry Nguyen.

PRESENTATIONS

Oral paper presentation, KDD 2025   Toronto, Canada

ACHIEVEMENTS

- Second Prize in Automatic Speech Recognition - Task 1 of VLSP 2022 competition (*Certificate*).
- Third Prize in Automatic Speech Recognition - Task 2 of VLSP 2022 competition (*Certificate*).
- Third Prize in Scientific Research Contest for College Students (*Certificate*).
- UET Merit Scholarship (8 semesters) - VNU University of Engineering and Technology.

TECHNICAL SKILLS

Programming	Python, Bash, C/C++, LaTeX, Java, SQL, JavaScript, HTML/CSS.
Libraries	PyTorch, TensorFlow, Scikit-learn, TensorFlow Serving.
Developer Tools	Jupyter, Git, Linux, Docker, MongoDB.

LANGUAGES

Vietnamese	Native
English	Advanced (IELTS 7.5)

REFERENCES

- **Prof. Truong-Son Hy**, Tenure-Track Assistant Professor, The University of Alabama at Birmingham, United States
  Email: thy@uab.edu. Homepage: https://hytruongson.github.io/HySonLab
- **Dr. Van Nguyen**, Head of Research, FPT Software, United States.
  Email: vannth19@fpt.com. Homepage: https://www.linkedin.com/in/vannguyen8/
- **Dr. Long Quoc Tran**, Senior Lecturer (eq. Assistant Professor), VNU University of Engineering and Technology, Vietnam
  Email: tqlong@vnu.edu.vn. Homepage: https://scholar.google.com/citations?user=tnDavuYAAAAJ