

# 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.
- Investigated 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*

- Engineered and open-sourced an [error analysis toolkit](#) for Multi-Target Multi-Camera tracking (MTMCT) systems.
- Constructed a Graph Neural Network architecture to address the re-identification challenge in MTMCT systems.
- 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 ( $\approx 70\text{ms}$  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

*GPA: 3.71/4.0 — Ranked 10th/153 (Top 7%) of the Computer Science Department*

## 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).

## Under Submission

### 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.

### 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.

### 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.

### AgentSound: A LLM-Driven Multi-Agent Framework for Augmenting Sounds with Semantic Context Guidance

Van Truong Hoang, Viet Pham, **Thanh V. T. Tran**, Van Nguyen, Harry Nguyen.

## PRESENTATIONS

Oral paper presentation, KDD 2025

Toronto, Canada

## ACHIEVEMENTS

(\*) VLSP is one of Vietnam's most prestigious NLP competitions, recently indexed as an ACL workshop.

- 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 (6 semesters) - VNU University of Engineering and Technology.

## TECHNICAL SKILLS

<b>Programming</b>	Python, Bash, C/C++, $\text{\LaTeX}$ , Java, SQL, JavaScript, HTML/CSS.
<b>Libraries</b>	PyTorch, TensorFlow, Scikit-learn, TensorFlow Serving.
<b>Developer Tools</b>	Git, Linux, Docker, MongoDB.

## LANGUAGES

<b>Vietnamese</b>	Native
<b>English</b>	Advanced (IELTS 7.5)

## ACTIVITIES

### Data Science Mentor

Jul 2023 - Aug 2023

Math and Science Summer Program (MaSSP)

Hanoi, Vietnam

- Taught Calculus and Computer Vision to high-school students.
- Supervised a group of 4 students on hand gesture recognition project, covering data processing, modeling, evaluation, and presentation.

## REFERENCES

- **Prof. Truong-Son Hy**, Tenure-Track Assistant Professor, The University of Alabama at Birmingham, United States  
Email: [thy@uab.edu](mailto:thy@uab.edu). Homepage: <https://hytruongson.github.io/HySonLab>
- **Dr. Van Nguyen**, Head of Research, FPT Software, United States.  
Email: [vannth19@fpt.com](mailto:vannth19@fpt.com). Homepage: <https://www.linkedin.com/in/vannnguyen8/>
- **Dr. Long Quoc Tran**, Senior Lecturer (eq. Assistant Professor), VNU University of Engineering and Technology, Vietnam  
Email: [tqlong@vnu.edu.vn](mailto:tqlong@vnu.edu.vn). Homepage: <https://scholar.google.com/citations?user=tnDavuYAAAAJ>