

Tran Quang Chung, PhD Student

✉ tran.quang_chung.tq9@naist.ac.jp bktranquangchung@gmail.com

🌐 <https://www.linkedin.com/in/chungtq/>

📞 Vietnam: 0965957672 Japan: 08035703887

🌐 **Personal Website:** <https://tranquangchung.github.io>

Education

- 2022 – 2025 📖 **Ph.D Student, 1.5 year at JAIST + 1.5 year at NAIST**
Thesis title: *Multimodal Speech Synthesis: Leveraging Diverse Sources of Text, Images.*
- 2019 – 2021 📖 **M.Sc. Computer Science, Hanoi University of Science and Technology**
Thesis title: *An Improved UNets Architecture and Its Applications.*
- 2014 – 2019 📖 **B.Sc. Computer Science, Hanoi University of Science and Technology**

Employment History

- 2023-Now 📖 **PhD AI Research** at NICT (Japan).
- 2019-2022 📖 **AI Research** at Computer Science Lab - Hanoi University of Science and Technology (HUST-Vietnam).
- 2018-2023 📖 **AI Engineer** at Vietnam Artificial Intelligence Solutions (Vietnam).

Research Publications (Selected)

Journal Articles

- 1 **Chung Tran**, Mai, L. C., & Sakriani, S. (2025). Zero-shot cross-lingual text-to-speech with style-enhanced normalization and auditory feedback training mechanism. *IEEE Transactions on Audio, Speech and Language Processing*, 33, 1479–1492. 🔗 doi:10.1109/TASLPRO.2025.3548429
- 2 Na, I. S., **Chung Tran**, Nguyen, D., & Dinh, S. (2020). Facial UV map completion for pose-invariant face recognition: A novel adversarial approach based on coupled attention residual unets. *Hum. centric Comput. Inf. Sci.*, 10, 45. 🔗 doi:10.1186/s13673-020-00250-w

Conference Proceedings

- 1 **Chung Tran**, & Sakriani Sakti. (2025). From Pixels to Voice: A Simple and Efficient End-to-End Spoken Image Description Approach via Vision Codec Language Models. In *ICASSP 2025*.
- 2 Ahmad Alfani Handoyo, **Chung Tran**, Dessi Puji Lestari, & Sakriani Sakti. (2024). Indonesian-English Code-Switching Speech Synthesizer Utilizing Multilingual STEN-TTS and BERT LID. In *O-COCOSDA 2024*.
- 3 **Chung Tran**, Luong, C. M., & Sakti, S. (2023). STEN-TTS: Improving Zero-shot Cross-Lingual Transfer for Multi-Lingual TTS with Style-Enhanced Normalization Diffusion Framework. In *INTERSPEECH 2023*.
- 4 **Chung Tran**, Huyen, H. C., & Sang, D. V. (2020). A Novel Generative Model to Synthesize Face Images for Pose-invariant Face Recognition. In *MAPR 2020*.
- 5 Sang, D. V., **Chung Tran**, Dung, N. D., & Na, I. S. (2020). Attention ResCUNet-GAN: A Novel Facial UV Map Completion for Pose-invariant Face Recognition. In *HCIS 2020*.

Research/Work Experience

Project Name	Year	Description
Automatic speech recognition (ASR)	2020-now	Model: Wav2Vec 2.0 Tool: SpeechBrain Role: Collect data, implement and deploy the model
Text-to-Speech	2020-now	Model: FastSpeech2, VITS, YourTTS, STEN-TTS Role: collect data, implement and deploy the model Publication: Interspeech2023
Speaker Recognition	2020-2021	Model: Thin Resnet-34 and Ecapa-TDNN Role: Finetune for Vietnamese voice Publication: RIVF-2021
Face Recognition and Recognition	2019-2022	Model: Attention ResCUNET Role: Implement and propose new ideas Publication: MAPR-2020, HCIS 2020, SCI-Journal Q1
Image-to-Speech	2022-now	Role: Propose new models to improve the performance Publication: ICASSP 2025

Research Topic

📌 Deep learning, Machine learning, Computer Vision, Natural Language Processing, Speech Processing.

Skills

Coding 📌 Python, C/C++, NodeJS, \LaTeX , ...
Databases 📌 MySQL, PostgreSQL
Systems 📌 Linux Operating System, Docker, Microservice
Misc. 📌 Academic research, teaching, training, consultation.

Miscellaneous Experience

Awards and Achievements

Aug 2022 📌 **The prestigious Japanese Government (MEXT) Scholarship** for PhD Student at JAIST.
Nov 2021 📌 Runner-up in the hackathon organized by AngelHack
Dec 2020 📌 Third Rank for VLSP Competition for the machine translation task.
Nov 2019 📌 **Prestigious prize in the field of Information Technology** from the Vietnam Talent Program.
Jan 2019 📌 Top1 in VietAI Machine Learning Class.
Nov 2018 📌 Runner-up of Cinnamon Intelligence Program.