Tran Quang Chung, PhD Student

tran.quang_chung.tq9@naist.ac.jp chungtq@jaist.ac.jp bktranquangchung@gmail.com
https://www.linkedin.com/in/chung-tran-625433118/
Vietnam: 0965957672 Japan: 08035703887
Personal Website: https://tranquangchung.github.io

Education

Apr 2024 – 2025 **Ph.D Studen, Nara Institute of Science and Technology (NAIST)**
Thesis title: *Multimodal Speech Synthesis: Leveraging Diverse Sources of Text, Images, and Brain Signals.*

2022 – Mar 2024 **Ph.D Studen, Japan Advanced Institute of Science and Technology (JAIST)**
Thesis title: *Multimodal Speech Synthesis: Leveraging Diverse Sources of Text, Images, and Brain Signals.*

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**
Thesis title: *Stock price prediction based on deep neural networks.*

Employment History

2021-2022 **AI Leader and Research** at Rikkei AI.

2019-now **AI Research** at Computer Science Lab - Hanoi University of Science and Technology (HUST).

2018-2020 **AI Leader and AI Engineer** at Vietnam Artificial Intelligence Solutions (VAIS).

Research Publications

Journal Articles

1. Ngoc, P. P., **Quang, Chung Tran**, & Chi, M. L. (0). Improving few-shot multi-speaker text-to-speech adaptive-based with extracting mel-vector (emv) for vietnamese. *International Journal of Asian Language Processing*, *0*(ja), null. doi:10.1142/S2717554523500042. eprint: https://doi.org/10.1142/S2717554523500042

2. Ngoc, P. P., **Quang, Chung Tran**, & Chi, M. L. (2023). Adapt-tts: High-quality zero-shot multi-speaker text-to-speech adaptive-based for vietnamese. *Journal of Computer Science and Cybernetics*, *39*(2), 159–173. Retrieved from https://vjs.ac.vn/index.php/jcc/article/view/18136

3. 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**, Luong, C. M., & Sakti, S. (2024). Maintaining Personal Styles in Multilingual TTS with STEN Approach in Diffusion Framework. In *Asj 2024*.

2. **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 *Proc. interspeech 2023* (pp. 4464–4468). doi:10.21437/Interspeech.2023-2243

3. Phuong Pham Ngoc, **Chung Tran**, & Mai Luong Chi. (2022). Improving a few-shot multi-speaker text-to-speech adaptative-based with extracting mel-vector (emv) for vietnamese. In *The 25th conference of the o-cocosda*.

4. **Chung Tran**, Quang Minh Nguyen, Phuong Pham Ngoc, & Quoc Truong Do. (2021). Improving speaker verification in noisy environment using dnn classifier. In *The 15th ieee-rivf international conference on computing and communication technologies* (pp. 1–6).

5. Phuong Pham Ngoc, **Chung Tran**, Truong Do Quoc, & Mai Luong Chi. (2021). A study on neural-network-based text-to-speech adaptation techniques for vietnamese. In *The 24th conference of the oriental cocosda*.

6. Ngoc Phuong Pham, **Chung Tran**, Nguyen Quang Minh, & Do Quoc Truong. (2020, December). Improving prosodic phrasing of Vietnamese text-to-speech systems. In *Proceedings of the 7th international workshop on vietnamese language and speech processing* (pp. 19–23). Hanoi, Vietnam: Association for Computational Linguistics. Retrieved from https://www.aclweb.org/anthology/2020.vlsp-1.4

7. **Chung Tran**, Huyen, H. C., & Sang, D. V. (2020). A novel generative model to synthesize face images for pose-invariant face recognition. In *2020 international conference on multimedia analysis and pattern recognition (mapr)* (pp. 1–6). doi:10.1109/MAPR49794.2020.9237763

8. 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 workshop 2020*.

Research Experience

Project Name	Year	Description
Vietnamese ASR (Automatic speech recognition)	2020-now	Model: HMM-DNN, 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, CSC Journal 2023, Internation Journal ALP 2023, Cocosda-2021, VLSP 2020**
Speaker Recognition	2020-2021	Model: Thin Resnet-34 and Ecapa-TDNN Role: Finetune for Vietnamese voice **Publication: RIVF-2021**
Vietnamese NMT (Neural machine translation)	2020-2021	Model: Transformer Tool: Tensor2Tensor Role Collect data (Vietnamese, Chinese, English) Implement and deploy the model

Research Experience

Project Name	Year	Description
Face Recognition and Recognition	2019-now	Model: Attention ResCUNET Role: Implement and propose new ideas **Publication: MAPR-2020, HCIS 2020, SCI-Journal Q1**
Polyp Segmentation	2020-now	Model: Propose the new methods/ideas https://github.com/tranquangchung/AG-CUResNeSt
Image Captioning	2021-2022	Model: Show Attend and Tell
Image to Speech	2022-now	Role: Propose new models to improve the performance

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	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	The first prize in information technology field for Vietnam Talent Program.
Jan 2019	Top1 in VietAI Machine Learning Class.
Nov 2018	Runner-up of Cinnamon Intelligence Program.