Huu Kim Nguyen

(+82) 10-5929-7054 | huukim136@gmail.com | linkedin.com/in/kim-nguyen136 | https://avocado136.github.io/

Interests

Deep Learning, Speech and Audio Signal Processing

Speech synthesis, voice conversion, generative models

Education

Yonsei University - top 3 universities in Korea — Seoul, Korea

M.S. in Electrical and Electronic Engineering — *Sept. 2019 – Sept. 2021*

- Digital Signal Processing & Artificial Intelligence (DSP&AI) Lab. (Prof. Hong-Goo Kang)
- Major: Speech signal processing, Deep learning

Hanoi University of Science and Technology — Hanoi, Vietnam

B.S. in Electronic and Telecommunication Engineering — *Sept. 2013 – Aug. 2018*

- Signal Processing and Radio Communication (SPARC) Lab. (Prof. Huy-Dzung Han)
- Major: Hardware design, FPGA design, Internet of Things, Machine learning

Experience

Data Scientist — Sept. 2021 – Present

LOVO.AI

- Develop neural speech synthesis models that fit customers' needs
- Train transformers-based autoregressive and non-autoregressive speech generative models
- Develop **LLM-based text-to-speech** models for both English and Korean languages
- Develop a **voice cloning** framework that is realistic and **human-like** level.
- Provide solutions for speaker generation from a limited dataset with limited speakers
- Develop emotional text-to-speech models with **natural prosody**
- Build voice conversion to convert speech to a target speaker's timbre
- Participate in **data collection and processing** pipeline
- Improve **pronunciation accuracy** of **BPE token-based Transformer-based TTS** models via cross attention
- Diffusion distillation for **faster inference speed**

Graduate Researcher — Sept. 2019 – Aug. 2021

DSP&AI lab - Yonsei University — *Seoul, Korea*

- Research speech-related topics e.g. speech synthesis, voice conversion
- Research solutions for speech synthesis on-device applications

Undergraduate Research Assistant — Jan 2017 – Feb. 2019

SPARC lab - Hanoi University of Science and Technology — *Hanoi, Vietnam*

- Build a smart algae cultivation system based on IoT platform
- Develop a secure remote FPGA reconfiguration method while the device is in operation

Projects

Genny (former Voicelab) - https://genny.lovo.ai/ — LOVO.AI

Create emotional TTS models for audio content creation. Research and develop LLM-based models for realistic voices

- Include +30 types of emotions and styles
- Focus on improving **natural prosody** and **high fidelity**
- Optimize the cost of autoregressive **LLM-based** models
- Tackle the known stability issues in autoregressive TTS models: **word skipping**, **repetition**, **babbling**, etc.

Voice Cloning — LOVO.AI

*Develop a **voice cloning** framework that is realistic and **human-like** level*

- Make use of powerful architecture of LLMs
- Improve voice cloning capability **drastically** via a unique prompting method

Voice Conversion LOVO.AI

Build a fast and low-latency any-to-many voice conversion model

- Tackle with disentanglement issue of timbre, pitch, linguistic information, energy
- Improve model's generalization for unseen combination of timbre, pitch, linguistic information, energy
- Develop accurate pitch modelling to preserve target speaker's pitch patterns

Development of Attribute Controllable Natural Keyword Speech Generation Method Qualcomm Korea

Speech augmentation in preparation for automatic speech recognition Nov. 2019 – Jun. 2020

- Research non-parallel voice conversion to synthesize speech utterances

Real-time Neural Text-to-speech on CPU Device Naver Corp.

Effective text-to-speech model for on-device applications Oct. 2020 – Aug. 2021

- Design a small-sized, fast-synthesizing text-to-speech model for portable devices
- Research non-autoregressive Transformer-based speech synthesis

VoiceVerse NFTs project - https://www.voiceverse.com/ LOVO.AI

Create a multi-speaker text-to-speech model with 8888 unique artificial voices

- Each voice is minted as an NFT token
- Owner of the voice token can use the accompanied text-to-speech tool

Publications

[1] Thi-Thai Yen Doan, Minh-Tri Ho, **Huu-Kim Nguyen**, and Huy-Dung Han. "Optimization of Spirulina sp. Cultivation using Reinforcement Learning with State Prediction based on LSTM Neural Network". In: *Journal of Applied Phycology* (2021).

[2] Kihyuk Jeong, **Huu-Kim Nguyen**, and Hong-Goo Kang. "A Light and Fast Text-To-Speech Model with Spectrum and Waveform Alignment Algorithms". In: *Proc. EUSIPCO*. 2021.

[3] **Huu-Kim Nguyen**, Kihyuk Jeong, and Hong-Goo Kang. "A Fast and Lightweight Speech Synthesis Model based on FastSpeech2". In: *Proc. ITC-CSCC*. 2021.

[4] **Huu-Kim Nguyen**, Kihyuk Jeong, Seyun Um, Min-Jae Hwang, Eunwoo Song, and Hong-Goo Kang. "LiteTTS: A Lightweight Mel-spectrogram-free Text-to-wave Synthesizer Based on Generative Adversarial Networks". In: *Proc. INTERSPEECH*. 2021.

Skills

Programming Languages and Frameworks: Python, Pytorch, Latex
Developer Tools: Git, Docker, Vim
Libraries: NumPy, pandas, Matplotlib, librosa, gradio, Hydra, multiprocessing, transformers, praat.
Languages: Vietnamese (native), English (professional working proficiency), Korean (elementary proficiency)