Nhan Tri-Do

Speech Research Engineer

☎ (+84) 0375-99-4743 | ✉ dotrinhan99@gmail.com | ⌨ v-nhandt21 | in dtnhan

Education

Advanced Program in Computer Science, *VNUHCM - University of Science*

GPA: 3.83/4.0

Topic: Speech Synthesis, Voice Conversion, Voice Cloning, Anti-Spoofing, Speaker Diarization

Experience

Vingroup Big Data Institute - Vingroup JSC   Oct 2021 - Present

AI Engineer   Hanoi, Vietnam

- Researched the Antispoofing problem involving Logical and Physical attack. Designed a model architecture that combines Rawboost Augmentation and Feature Extractor, a speech self-supervised learning model, with a Graph Attention Neural Network to detect spoof audio (2024).

- Designed the Smart AI Voice Recording service for insurance operations, aimed at protecting users by verifying contract transparency through Speaker Diarization, Speech Recognition, and identifying agents using Voice Biometric to prevent fraud. (2024)

- Implemented quality control workflows for the Enrollment and Verification modules in the Government's Voice Biometric project. Developed a gateway to coordinate microservice modules, including SNR Estimator, VAD, ASR, and Speaker Counter, packaged as TorchServe and connected to a Vector Database, Milvus. (2023)

- Optimized GPU resource usage for the Speech Synthesis Service by 4x by developing a Multispeaker Acoustic model to replace all individual voice models. Additionally, the model can adapt to new voices with just 30 minutes of studio-quality data and perform zero-shot voice cloning with 30 seconds of voice data through an automated processing and feature extraction framework. (2023)

- Optimized the Vocoder model for the Cloud Speech Synthesis module in Electric Vehicle Assistant, by implementing Universal Multistream Hifigan with Conditional Speaker to achieve a 1.5x faster inference time in locust load test while preserving voice quality. (2022)

- Developed User Interfaces for AI services to support demos, data labeling, model evaluation, POC, and project bidding, primarily using Streamlit, Material UI React, and Wavesurfer.js (2022)

- Enhanced the robustness of the Autoregressive Acoustic model - Tacotron2 in the Speech Synthesis module by integrating Monotonic Alignment Attention, resulting in better alignment between phonemes and speech, more distinct voice output, and elimination of uncontrolled sounds (2021)

- Designed and developed a Massive Speech Data Crawler, an automated data crawling and processing system. The system includes automatic data selection, signal formatting, vocal removal, data segmentation, denoising, and speech enhancement, transcription, quality ranking, and automatic data characteristic analysis (2021)

VinAI Research Institute - Vingroup JSC   Dec 2020 - Sep 2021

Resident Engineer   Hanoi, Vietnam

- Integrated a state-of-the-art Text-to-Speech AI model for Vietnamese Voice Control in electric vehicles, using Mel-Generator (Tacotron2, Fastspeech2, GlowTTS) and Vocoder (Multiband-Melgan, Hifigan), and optimized it with ONNX Runtime for edge device inference

- Designed and deployed SpeechMT, a real-time voice-based machine translation service, with each service containerized using Docker and interfacing through FastAPI
- Deployed a Speech-based QA Car Manual, a lightweight system that interacts with users via speech commands

KMS Technology - Outsourcing and Consulting Enterprise
June 2020 – Sep 2020

Software Engineer   Ho Chi Minh city, Vietnam

- Researched core AI for RASA-based chatbot platform, improving the accuracy of the Intent Classifier (DIET) on SQUAD data by combining Dense (ConveRT and BERT) and Sparse (TF-IDF and Count Vector) features.
- Developed backend services for Reading Comprehension by training the Bi-Directional Attention Flow (BiDAF) model using AllenNLP, managed multiple requests and service deployment using Docker Compose, Swagger, Flask, and RabbitMQ for message queue handling limited resources.

Scooter Saigon Tour – Tourism Startup
Jan 2014 – Aug 2017

SEO - Developer   Ho Chi Minh city, Vietnam

- Developed a website using WordPress, set up the PHP backend, and integrated a payment system with VTCpay
- Performed Search Engine Optimization (SEO) by optimizing website content and building a network of back-links and social media connections to enhance website rankings

Freelance Side Jobs .

- Voice-Preserving Speech Machine Translation
- Speech Enhancement and Noise Reduction
- Singing Voice Conversion
- Sleep Stage Classification

Lab Research

Speech and Language lab - Nanyang Technological University
Sep 2022 – Mar 2023

Internship   Remote Singapore

- Participated in a government project with ST Engineering, developing and testing a model for classifying emergency sound events in urban areas.

Artificial Intelligence lab - University of Science
Feb 2019 – Sep 2021

Collaborator   Ho Chi Minh city, Vietnam

- Assistant at VNsigma – class for python beginner, supported by American Consulate.
- Participated in organizing ERC2019 – Emotion Recognition Challenge
- Built a Vietnamese speech corpus (19GB) and conducted research on Text Normalization and Phonology.

Robotics and IoT lab - University of Science
May 2019 – Sep 2021

Teacher assistant   Ho Chi Minh city, Vietnam

- Teacher for EV3 Mindstorm and Coach for First Lego League (FLL) and World Robot Olympiad (WRO).

Computational linguistics center - University of Science
April 2019 – Sep 2019

Teacher assistant   Ho Chi Minh city, Vietnam

- Teacher Assistant for Faculty of English linguistics, USSH in SDL Trados Studio - computer-assisted translation software.

Publications

Voice Conversion - Voice Cloning

- Thesis: DeepSpeechVC: Voice Cloning Framework with Speech Synthesis and Voice Conversion - Experiment with Speech2Speech techniques to make voice conversion from a small sample of the target speakers

  Tri-Nhan Do, Minh-Tri Nguyen under the advice of Prof. Vu Hai Quan, Msc. Xuan-Nam Cao

Speech Synthesis

- Journal: FastSpeechStyle: Fast, Emotion Controllable and High-Quality Speech Synthesis at The International Journal of Asian Language Processing

  Van Thinh Nguyen, **Tri-Nhan Do**, Hung-Cuong Pham, Tuan Vu Ho, Ngoc-Minh-Khanh Nguyen, Dang-Khoa Mac

- Paper: Vietnamese Speech-based Question Answering over Car Manuals - Data-Centric AI Workshop at 35th Conference on Neural Information Processing Systems (DCAI@NeurIPS2021)

  Tin Duy Vo, Manh Tien Luong, Duong Minh Le, Hieu Minh Tran, **Nhan Tri Do**, Tuan-Duy Hien Nguyen, Hung Hai Bui, Dat Quoc Nguyen, Dinh Quoc Phung

- Paper: Vietnamese Speech Synthesis with End-to-End Model and Text Normalization - 2020 7th NAFOSTED

  Do Tri Nhan, Nguyen Minh Tri, Cao Xuan Nam

- Pypi Package Vinorm: Vietnamese Normalization that converts raw text to spoken form with three version C++, NDK for Android and PyPi - 12,739 downloads

- Pypi Package Viphoneme: Phonetization that converts Vietnamese Grapheme to IPA - 1112 downloads

Speech and Signal Classification

- Paper: Emotion Classification Using Wavenet Features with SpecAugment and EfficientNet - MediaEval'20, December 14-15 2020, France

  Tri-Nhan Do, Minh-Tri Nguyen, Hai-Dang Nguyen, Minh-Triet Tran, Xuan-Nam Cao

- Paper: Sound Event Detection with Soft Labels Using Self-Attention Mechanisms for Global Scene Feature Extraction - Detection and Classification of Acoustic Scenes and Events 2023 (DCASE)

  Nhan Tri-Do, Param Biyani, Zhang Yuxuan, Andrew Koh Jin Jie, Chng Eng Siong

Awards

- Outstanding Employee of the Year 2023 at VinBigdata   2023
- First Place at VLSP TTS: Emotional speech synthesis   2022
- Outstanding Student Award in the Field of Artificial Intelligence Ho Chi Minh City   2021
- Excellent Thesis - Science and Technology Student Award   2021
- First Prize Science-A-Thon Award in 9th Vietnam Summer School of Science   2022
- Runner Up in KMS Hackathon   2020

Certifications and Courses

- Contributed to obtaining ISO/IEC 19795-1:2021 under NIST/NVLAP criteria for Voice Biometric Testing.   2024
- Participation in IEEE Spoken Language Technology Workshop Hackathon   2022
- 7th-place / 23 teams at The IEEE Signal Processing Society's Signal Processing Cup   2022
- FPT Scholarship for Final Round in Code War Competition   2019
- Final Round Samsung Collegiate Programming Cup   2018
- ACM-ICPC Vietnam National Programming Contest: rank 33   2018
- Self-Driving Cars Specialization course offered by University of Toronto on Coursera

1