

Do Tri Nhan

✉ dotrinhan99@gmail.com | 📞 +84-375-99-4743 | 🌐 v-nhandt21.github.io | 📍 Ho Chi Minh City, Vietnam

TECHNICAL SKILLS

Speech & Audio: TTS, Speaker Verification & Diarization, Voice Cloning & Conversion, Anti-spoofing, Speech Enhancement, Sound Event Detection

ML & LLM: PyTorch, ONNX, Triton, TorchServe, SLM, Reinforcement Learning, Agno Agent, RAG, Speech SSL

Engineering: Docker, FastAPI, Flask, RabbitMQ, Milvus, Qdrant Vector DB, Slurm, Datadog, Jenkins, Jfrog

Languages: Python (primary), C++, Java, React Vibe

EDUCATION

VNUHCM – University of Science

Bachelor of Science, Advanced Program in Computer Science (APCS K17)

Sep 2017 – Jun 2021
Ho Chi Minh City, Vietnam

- GPA: 3.83 / 4.0 TOEFL ITP: 550 / 677

EXPERIENCE

Vinfast JSC

AI Team Lead – Speech Generation

Jun 2026 – Present
Ho Chi Minh City, Vietnam

- **Unified TTS Serving Gateway:** Architected a production FastAPI gateway consolidating multiple TTS model tiers and normalization backends — supporting streaming, async job queue, API key auth with per-key usage tracking, and horizontal shard deployment for multilingual production serving.
- **TTS Normalization Automation Harness:** Designed an end-to-end QC automation framework with LLM-based error classification, BA-configurable per-project switch-code vocabulary management, automated Java SDK and prompt rebuild, and full regression evaluation — reducing engineering intervention for vocabulary updates to zero.

VinSmart Future – Vingroup JSC

AI Team Lead – Speech Generation

Feb 2026 – Jun 2026
Ho Chi Minh City, Vietnam

- **Research Team Leadership:** Lead a team of 6 members on speech generation: voice design, non-verbal audio, and reinforcement learning for speech synthesis.
- **Voice Model Product Management:** Coordinate between research team and product projects; standardize quality assurance and release processes for voice models across three product lines: robot assistants, in-vehicle AI, and super-app voice mode.

Senior AI Engineer – Core AI Research

Jul 2025 – Feb 2026

- **Vietnamese Phoneme-Based Flow Matching TTS:** Developed a Vietnamese phoneme-adaptive fine-tuning TTS model achieving stable quality with low WER in production.
- **TensorRT & Triton Serving Optimization:** Optimized model serving with TensorRT and Triton Inference Server, achieving 2× throughput improvement and increased concurrent user capacity.
- **Hybrid Text Normalization:** Developed a Small Language Model (SLM) combined with rule-based normalization in a hybrid pipeline, improving accuracy by 13% over the rule-based baseline.

Vingroup Big Data Institute – Vingroup JSC

Middle AI Engineer – AI Agent & Voice Biometric

Jan 2024 – Jul 2025
Hanoi, Vietnam

- **AI Agent with Knowledge Base:** Built a RAG-based AI agent with query processing, embedding, function calling, and automatic double-bot quality control evaluation (2025).
- **Speech Data Warehouse:** Managed a team building a speech data warehouse for storage, querying, processing, release management, and natural language querying integration - Data Intelligence (2025).
- **Streaming Diarization Optimization:** Optimized and served streaming diarization with Triton EEND-VC, achieving 2× speed and superior quality over third-party solutions (2025).
- **Anti-spoofing Module:** Researched and deployed an anti-spoofing system with an EER of 3.75%, combining RawBoost augmentation, speech SSL, and Graph Neural Networks (2024).

AI Engineer – Speech Synthesis & Voice Biometric

Oct 2021 – Dec 2023

- **Smart AI Voice Recording Service:** Designed an AI voice recording service for insurance operations using Speaker Diarization, ASR, and Voice Biometrics to detect agent fraud (2023).
- **Government Voice Biometric Project:** Implemented QC workflows and a microservice gateway contributing SNR Estimator, VAD, ASR, and Speaker Counter via TorchServe and Milvus (2023).
- **Multispeaker Acoustic Model:** Reduced GPU usage 4× via a multispeaker acoustic model supporting 30-min fine-tuning and zero-shot cloning with 30 s of audio (2023).
- **Universal Multistream Vocoder:** Optimized the EV assistant cloud TTS vocoder with Universal Multistream HiFi-GAN, achieving 1.5× faster inference without quality loss (2022).
- **AI Service User Interfaces:** Developed UI tools for AI services to support demos, data labeling, model evaluation, POC, and project bidding using Streamlit, Material UI React, and Wavesurfer.js (2022).

- **Tacotron2 Enhancement:** Enhanced robustness of Tacotron2 by integrating Monotonic Alignment Attention, improving phoneme–speech alignment and eliminating noise artifacts (2021).
- **Massive Speech Data Crawler:** Designed a large-scale automated speech data crawling and processing system covering data selection, formatting, denoising, transcription, quality ranking, and characteristic analysis (2021).

VinAI Research Institute – Vingroup JSC

Engineering Resident – ERP Batch 1

Dec 2020 – Sep 2021

Hanoi, Vietnam

- **Vietnamese TTS for Electric Vehicles:** Integrated state-of-the-art TTS (Tacotron2, FastSpeech2, GlowTTS, HiFi-GAN), optimized with ONNX Runtime for edge inference.
- **SpeechMT:** Designed and deployed a real-time voice-based machine translation service containerized with Docker and served via FastAPI.
- **Speech-Based QA for Car Manuals:** Developed a lightweight speech-interactive system for voice-driven car manual queries, deployed end-to-end.

KMS Technology

Software Engineer – Center of Excellence

Jun 2020 – Sep 2020

Ho Chi Minh City, Vietnam

- **RASA Chatbot Core Research:** Improved intent classification by combining Dense (ConveRT, BERT) and Sparse (TF-IDF, Count Vectors) features; evaluated on SQuAD data.
- **Reading Comprehension Backend:** Built BIDAF-based reading comprehension services with AllenNLP, Flask, Swagger, Docker Compose, and RabbitMQ.

Scooter Saigon Tour – Tourism Startup

SEO Developer – Founder

Jan 2014 – Aug 2017

Ho Chi Minh City, Vietnam

- **Website Development:** Built a WordPress website with PHP backend and integrated VTCpay payment system.
- **SEO & Online Marketing:** Optimized site content, built backlink networks, and managed social media channels to improve search rankings.

Freelance AI Projects (selected):

- Voice-Preserving Speech MT
- Singing Voice Conversion
- Pronunciation Assessment
- Speech Enhancement & Noise Reduction
- Sleep Stage Classification
- Emotional Dubbing

RESEARCH AFFILIATIONS

Speech and Language Lab, Nanyang Technological University

Research Intern (Remote)

Sep 2022 – Mar 2023

Singapore

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

Artificial Intelligence Lab, University of Science

Research Assistant

Feb 2019 – Sep 2021

Ho Chi Minh City, Vietnam

- Built a Vietnamese speech corpus (19 GB) and conducted research on text normalization and phonology.
- Organized ERC2019 – Emotion Recognition Challenge.
- Teaching assistant at VNSigma Python beginner class, supported by the American Consulate.

Robotics & IoT Lab, University of Science

Research Assistant

May 2019 – Sep 2021

Ho Chi Minh City, Vietnam

- Teacher for EV3 Mindstorm and coach for First Lego League (FLL), EVJ Makethon, and World Robot Olympiad (WRO) competitions.

Computational Linguistics Center, University of Science

Research Assistant

Apr 2019 – Sep 2019

Ho Chi Minh City, Vietnam

- Teaching assistant for the Faculty of English Linguistics (USSH), guiding students in SDL Trados Studio, a computer-assisted translation software.

PUBLICATIONS

- Unified Acoustic Representation Learning for Vietnamese Speech Classification Tasks. *MAPR 2025*.
- Analyzing the Correlation and Impact of Speech Evaluation Metrics on Speaker Verification and ASR. *MAPR 2025*.
- Adapting WavLM for Vietnamese Speaker Diarization in Real-world Conversations. *MAPR 2025*.
- Enhancing Deepfake Detection: WavLM and Advanced RawBoost Augmentation. *RIVF 2024*.
- Voice Attacker: Multi-Head Factorized Attentive Reconstructor and Gradient Reversal for Random Prosody Anonymization. *Voice Privacy 2024*.
- Sound Event Detection with Soft Labels Using Self-Attention for Global Scene Feature Extraction. *DCASE 2023*.
- FastSpeechStyle: Fast, Emotion-Controllable, High-Quality Speech Synthesis. *International Journal of Asian Language Processing*.
- Vietnamese Speech-based Question Answering over Car Manuals. *DCAI/NeurIPS 2021*.
- Vietnamese Speech Synthesis with End-to-End Model and Text Normalization. *7th NAFOSTED 2020*.
- Emotion Classification Using WaveNet Features with SpecAugment and EfficientNet. *MediaEval 2020*.

PyPi package `vinorm`: Vietnamese Text Normalization (34,933+ downloads / 6 months).

PyPi package `viphoneme`: Vietnamese Grapheme-to-IPA Phonetization (4,195+ downloads / 6 months).

Thesis: DeepSpeechVC – Voice Cloning Framework with Speech Synthesis and Voice Conversion.

AWARDS & RECOGNITION

- Outstanding Employee of the Year – VinBigdata 2023
- First Place, VLSP TTS Emotional Speech Synthesis 2022
- First Prize, Science-A-Thon, 9th Vietnamese Summer School of Science 2022
- Outstanding Student Award in Artificial Intelligence, Ho Chi Minh City 2021
- Excellent Thesis – Science and Technology Student Award 2021
- Runner-Up, KMS Hackathon 2020

CERTIFICATIONS & COMPETITIONS

- Contributed to obtaining ISO/IEC 19795-1:2021 (NIST/NVLAP) for Voice Biometric Testing 2024
- Participation, IEEE Spoken Language Technology Workshop Hackathon 2022
- 7th/23 teams, IEEE Signal Processing Society's Signal Processing Cup 2022
- FPT Scholarship, Final Round, Code War Competition 2019
- Final Round, Samsung Collegiate Programming Cup 2018
- ACM-ICPC Vietnam National Programming Contest – Rank 33 2018
- Self-Driving Cars Specialisation – University of Toronto (Coursera) 2018

VOLUNTEERING

- Volunteer for medical research on dental surgery at HCMC Oromaxillofacial Hospital 2021
- F1 contact tracing support during the Covid-19 pandemic 2022
- Organized ERC2019 – Emotion Recognition Challenge 2019
- Teaching assistant at VNsigma Python beginner class (American Consulate) 2019