NGUYEN DINH CUONG

AI ENGINEER

+84 976 174 663 | dinhcuong1007@gmail.com | Hai Ba Trung, Hanoi, Vietnam | DOB: 10 Jul 2003

PROFESSIONAL SUMMARY

AI Engineer with hands-on experience delivering production Computer Vision, NLP, LLM, and Voice AI systems. Skilled in training, fine-tuning, and deploying models across OCR, face recognition/anti-spoofing, speech recognition, and text-to-speech pipelines. Aiming to leverage this expertise to drive product innovation while growing into an advanced AI specialist over the next two years, building efficient and creative technology solutions.

WORK EXPERIENCE

AI Engineer — InfinIQ | *Mar 2026 – Present*

AI Engineer — 3GTEL | *Jan 2024 – Present*

AI Engineer — REVVE | *Sep 2025 – Jun 2026*

Developed AI/ML projects across Computer Vision, NLP, LLM, and Voice domains; designed and built models for OCR, Face Recognition/Face ID, Text-to-Speech, Speech-to-Text, and Chatbot products.

Face Matching, Face ID & Face Anti-Spoofing

- **99.6% accuracy** in face recognition and similarity matching on a dataset of 67,000 images.
- Built an automated facial-recognition attendance feature integrated into an EKYC system, improving management efficiency for **200 employees**.

OCR Information Extraction System

- Built an ID-card (CCCD) OCR pipeline achieving **95–97% accuracy** on a dataset of 89,000 images.
- Integrated into an EKYC system supporting SIM-card registration, letting a single manager replace an entire data-entry team while reducing input errors.

Japanese Handwriting OCR Solution

- Developed a solution for extracting handwritten Japanese text from complex images, reaching **98% accuracy** on a dataset of 5,000 images.
- Reduced document-digitization staffing needs by **80%** (from 50 down to 2 employees).

Text-to-Speech (TTS) Systems

- **High-performance version:** optimized for ultra-low latency in automated call-center contact centers, ensuring seamless real-time interaction and reducing operational staffing needs.
- **High-fidelity version:** leveraged advanced Voice Cloning technology for web and multimedia production (podcasts, audiobooks), cutting voice-actor production costs.

Speech-to-Text (STT) System

- Optimized Vietnamese speech recognition models, achieving a Word Error Rate (WER) below **5–8%** on real-world general-domain data, with real-time performance for call centers and voicebot assistants.

Customer Support Chatbot

- Built a local chatbot integrating NLP + LLM + RAG + Agent, deployed city-wide in Hanoi, cutting support-team resolution time by **80%** for sales inquiries and bus route/schedule lookups.

SKILLS

AI/ML Expertise

- Computer Vision, Natural Language Processing (NLP), Deep Learning, Speech Processing (ASR/TTS).

Programming Languages

- Python, C, C++.

Frameworks & Libraries

- YOLO, PaddleOCR, PyTorch, OpenCV, FastAPI, LlamaIndex, LangChain, vLLM, Ollama, Agno, LiveKit.

Models Trained / Fine-tuned

- YOLO (object detection, anti-spoofing, license plate/head detection).
- PaddleOCR (text extraction, ID-card OCR, Japanese handwriting OCR).
- ZipVoice, VoxCPM, F5-TTS, OmniVoice (Text-to-Speech, Voice Cloning).
- Zipformer, Parakeet ASR, Nemotron-3.5-ASR-Streaming (Speech-to-Text / streaming ASR).

Tools

- Docker, Git, Label Studio.

Other Skills

- Image processing, system integration, multi-GPU Linux server operations.

EDUCATION & TRAINING

B.S., Artificial Intelligence & Data Science — Thuy Loi University (Oct 2021 – Sep 2025)

- Participated in research on chatbot model training and development.

AI Vietnam (May 2023 – May 2024)

- Completed intensive training programs in AI/ML, Computer Vision, and NLP.

ProtonX (Nov 2024 – Present)

- Ongoing coursework on AI Agents and language model training.

Big-O Coding (Feb 2022 – Apr 2022)

- Studied fundamental data structures and algorithms.

INTERESTS

- Passionate about researching new AI models and applying them to real-world problems.
- Strong self-learning ability and effective team collaboration.