Tran Dinh Manh Cuong

Ha Noi, Viet Nam cuong.tdm2000@gmail.com 0356780137 cuongtran2203 cuongtran2000
cuongtran2203

ABOUT ME

I am a highly motivated and experienced AI engineer with a strong academic background in computer science and machine learning. I have accumulated over three years of practical experience working on cutting-edge AI projects. My expertise includes developing and deploying machine learning models, and computer vision applications. I am passionate about staying up-to-date with the latest advancements in AI technology and am eager to contribute my skills and knowledge to drive innovation in the field. With a proven track record of delivering results, I am confident in my ability to excel in the role of AI engineer.

Education

BS **Ha Noi University of Science and Technology**-Mechatronics Engineer — Aug 2018 - Aug 2023

- Graduation project: Research and Design Segmentation model for navigation of quadruped robot
- Engineer: Research,Design optimal navigation strategy for autonomus robot based on deep learning technology and computer vision
- Co-author ISI(Q1) paper : IRDC-Net: Lightweight Semantic Segmentation Network Based on Monocular Camera for Mobile Robot Navigation
- Author paper : Application of Face Recognition Technique with ArcFace Model to Smart Camera System
- Author paper: Design of Ultra Fast Semantic Segmentation Model for Autonomous Mobile robot Navigation

Master Ha Noi University of Science and Technology-Mechatronics Engineer — Aug 2024 - Now

- Research and develop model AI for robot navigation

Award

- First prize in scientific research at school level
- Second prize in national scientific research
- Third prize in CodeWar

Experience

GMOZcom - Runsystem - 1 st Dao Duy Anh, Dong Da, AI Engineer — 02/2024 - Now

- Completed the tool for training the Stable Diffusion model: successfully finished building a training tool for the Stable Diffusion model.
- Successfully deployed an image generation application for Australian customers: successfully launched an application that generates images from prompts for clients in Australia.
- Researched and developed AI solutions: I was involved in researching and developing AI solutions at GMOZcom - Runsystem.
- Deployed models on torchserve and triton: I have experience deploying models on platforms like torchserve and triton.
- Participated in building a Japan license plate recognition application: I contributed to the project of developing an application for recognizing Japanese license plates.

AI Academy Viet Nam, AI Engineer — Ha Noi, Viet Nam
June 2022 – Feb 2024

- License plate recognition on highways with 95 percent accuracy and fast inference time (about 100ms).
- Participate in maintaining and developing many AI projects.
- Have experience deploying models on AI serving platforms such as torchserve and triton.

Publications

IRDC-Net: Lightweight Semantic Segmentation Network Based on Monocular Camera for Mobile Robot Navigation — 2023

Thai Viet-Dang, ***Tran Dinh Manh Cuong***, Phan Xuan Tan

https://doi.org/10.3390/s23156907

Projects

Synth Text — GMO - 02/2025 - Now

- Research and Develop model SynthText Japanese
- Preparing and processs data for Japanese text generator
- Generating Japanese text for training text recognition model
- Tools Used: Python, Stable Diffusion, Huggingface, Deeplearning

Information Extraction — GMO - 02/2025 - Now

- Research, using LLM for information extraction
- Research finetuning LLM technique
- Research LLM Training Strategy
- Research and deploy LLM serving (ollama,SGLang, vLLM,Llama.cpp)
- Tools Used: Python, Huggingface, unsloth, LLM serving

Image Generation — GMO - 02/2024 - 04/2024

- Develop Image Generation (Poster generation) application
- Tools Used: Python, Stable Diffusion, Huggingface

Japan License plate Recognition — GMO - 02/2024 - 04/2024

- Building app LPR with 4 output: hinagana character, issuing region, classification number, serial number
- Tools Used: Python, Computer Vision, TensorRT, C++

HAKARU AI — GMO - 04/2024 - Now

- The hakaru AI is clock ‘s information recognition software
- Analyze the information on the watch and give the necessary results
- Tools Used: Python, Flask, OpenCV, Torch, TensorFlow, Computer Vision

FACE ID — GMO - 08/2024 - Now

- Design and improve Face id system for the bank
- Research module Face Recognition and Parallel training solutions for big data
- Tools Used: Python, Triton Server, Computer Vision, Torchrun

FACE ID APP — AI Academy - 07/2022 - 12/2022

- The system included: Face Recognition and Face Liveness
- Solve the problem of automatic attendance
- Tools Used: Python, Computer Vision, Face Anti Spoofing, torchserve

Vehicle Detection System AI Academy - 07/2022 - 12/2022

- Contruct Vehicle Detection System include : counting car,vehicle classification,detect abnormal phenomena
- Tools Used: Python, Computer Vision, multi threading, Gstreaming, Deepstream app

ITS AI Academy - 06/2022 - 12/2022

- Recognize licesen plate for Highway with 95% acc on 1 core CPU, infer about 100ms
- Tools Used: Python, Computer Vision, OpenVino, Onnx

Height measurement AI Academy - 06/2022 - 12/2022

- Height mesurement based on kinnect camera and Deep learning technology
- Provide body parameters and assess health status
- Tools Used: Python, Computer Vision, Pose Estimation

Technologies

Languages: Python, C++

Framework AI: Tensorflow, pytorch, paddle, tensorRT, unsoth, huggingface, transformers, perf, ...

Computer vision package: Opencv,pillow,numpy,QT framework

AI serving: Triton Server, ollama, llamacpp, vLLM, SGLang

Others: MongoDB, docker, deepstream, Jetson nano board, ROS1