

NGUYEN LAM VY DAT

AI Engineer

 vydat1702@gmail.com  +84 888 644 887  Portfolio  Dat Nguyen-Lam Vy  VyDat-1702

SKILLS

Robot Learning: Imitation Learning, Behavior Cloning, Implicit Behavioral Cloning (IBC), Diffusion Policy, Flow Matching Policy, Vision-Language-Action Models (VLA), Reinforcement Learning

AI / ML: PyTorch, TensorFlow, Model Training - Evaluation - Optimization

Computer Vision: OpenCV, MediaPipe, Image Processing, Edge AI Deployment

Programming: Python, C++

Tools & Platforms: Linux, Git, GitHub, Docker, LaTeX, Jira

Languages: Vietnamese (Native), English (TOEIC 630/990)

EXPERIENCE

Research Engineer (Contract)

- Built an automated benchmarking pipeline across embedded devices, identifying critical bottlenecks to drive data-driven hardware and software optimizations.
- Executed real-world integration testing and authored technical reports, reducing deployment regressions and guiding product version upgrades.

SHAREWORK CORP
R&D Department
Ho Chi Minh City
Dec 2025 – Mar 2026

Research Engineer Intern

- Deployed TensorFlow Lite models on STM32 microcontrollers using quantization and pruning techniques, significantly reducing inference latency and energy consumption.
- Engineered an end-to-end IoT data pipeline (I2C/SPI to MQTT) and a real-time monitoring dashboard, ensuring reliable sensor acquisition and accelerating system validation.

SHAREWORK CORP
R&D Department
Ho Chi Minh City
Jun 2025 – Dec 2025

AI Engineer Intern

- Engineered Edge AI solutions for resource-constrained embedded systems, optimizing deep learning models to ensure efficient real-world deployment.
- Accelerated on-device inference performance by adapting complex network architectures for hardware-specific constraints.

Bosch Global Software Technologies
Ho Chi Minh City
Apr 2026 – Present

PUBLICATIONS

Hybrid Agentic Architecture for Vietnamese Sign Language: Orchestrating RAG and Syntactic Restructuring – **Dat Nguyen-Lam Vy**, Nghe-Nhan Truong, Van Dang Nguyen, Dao Tuan Dat Vo, and My-Ha Le

From Multi-View RGB-D to Single-View: Fusion and Distillation for Efficient Sign Language Recognition – Dao Tuan Dat Vo, Nghe-Nhan Truong, **Dat Nguyen-Lam Vy**, Van Dang Nguyen, and My-Ha Le

TKR-HMR: Tokenized Kinematic Representation for Human Mesh Recovery from Video – Van Dang Nguyen, Nghe-Nhan Truong, Dao Tuan Dat Vo, **Dat Nguyen-Lam Vy**, and My-Ha Le

MicroKWS-Net: An Ultra-Lightweight Attention-Guided CNN for Keyword Spotting on Edge Devices via Knowledge Distillation – Dao Tuan Dat Vo, Van Dang Nguyen, **Lam Vy Dat Nguyen**, et al., Minh Huan Vo

Accepted, in press

Accepted, in press

Accepted, in press

Accepted, in press

PROJECTS

Neural Kalman Filter for Lateral Velocity Estimation

Bosch Global Software
Technologies

- Built a Neural Kalman Filter with a TCN + Attention backbone to estimate vehicle lateral velocity (V_y) for the TI TDA4VM embedded platform, fusing the Ackermann formula and IMU integration via a learned Kalman gain $K[t]$; pretrained on Car-Maker simulation data and fine-tuned on real-world data using LoRA to close the sim-to-real gap.
- Reduced V_y estimation error vs. VSE (Ackermann-based baseline) by **48.6%** RMSE ($0.052 \rightarrow 0.027$ m/s) and **47.1%** MAE ($0.038 \rightarrow 0.020$ m/s), with the largest gains on low- μ surfaces and at high speeds where Ackermann-based estimation fails.
- Deployed the optimized model directly on TI TDA4VM via TIDLRT (EdgeAI TIDL tools).

Edge AI Module

SHAREWORK CORP

- Built an end-to-end Edge AI system converting raw sensor signals into frequency-spectrum images, with a custom CNN for automated machine health diagnosis and condition monitoring.
- Optimized inference using model quantization to reduce latency and resource usage on edge devices.
- Developed dashboards to interpret outputs, detect anomalies, and validate performance.

Robot Learning from Demonstration: IBC & Diffusion Policy

Personal Project
[IBC](#) 
[Diffusion Policy](#) 

- Reimplemented two imitation learning paradigms from scratch in PyTorch (50 demos/task) to study multimodal action modeling in low-data regimes.
- IBC (Block Push): energy-based policy trained via InfoNCE loss with DAgger-style interactive data aggregation; inference sampler combining DFO with Langevin dynamics + Gaussian noise for refinement. **88%** success rate (44/50) vs. 54% (27/50) for an MSE-regression BC baseline, against a 100% (50/50) oracle policy.
- Diffusion Policy (PushT): Transformer-based denoising diffusion action head under the same low-data regime, reaching **60%** success rate on this higher-precision task; identified diffusion step count and training horizon as the main levers for further improvement.

LICENSES & CERTIFICATIONS

PyTorch for Deep Learning Specialization

[View Credential](#) 

Google AI Essentials

[View Credential](#) 

Samsung Innovation Campus – Artificial Intelligence (AI)

[View Credential](#) 

Google Project Management Specialization

[View Credential](#) 

EDUCATION

BS **HCMC University of Technology and Education**
Embedded Systems and Internet of Things

2022 – 2026
GPA: 3.29/4.0