Arkar Phyo

AI Engineer

+66 910-681-379 | arkarphyo.akp1996@gmail.com | GitHub | LinkedIn | Website
Bangkok, Thailand

Summary

Machine Learning Engineer with over 5 years of experience designing and deploying full-stack AI solutions across computer vision, language models, and scalable backends. Proven expertise in Python-based AI engineering, GPU optimization, and cloud integration (AWS, Docker, Kubernetes). Passionate about building autonomous AI agents leveraging predictive and generative models. Strong communicator with global collaboration experience and business-level English and Japanese (JLPT N3).

Core Skills

Programming:	Python, C++, Typescripts
Fullstack:	React, Next,js, Vite, Qt, Flask, FastAPI
Machine Learning:	PyTorch, TensorFlow, HuggingFace, Transformers, YOLO, OpenCV, scikit-learn, Numpy, Matplotlib
MLOps & Cloud:	Docker, Kubernetes, Celery, RabbitMQ, Github Actions CI/CD, AWS
Databases:	MongoDB, Firebase, SQL
Other:	ROS2, TensorRT, ONNX, TFLite, Bazel

Education

Yangon Technological University (Dec 2012 - Oct 2018)
B.E. in Computer Engineering and Information Technology

Experience

Global Walkers Co., Ltd
Senior AI Engineer / Team Lead (2022 - 2025)   Yangon, Myanamar

- Designed and led scalable AI systems integrating object detection, tracking, and re-identification for smart city and retail applications.
- Developed multi-modal systems combining LLMs and VLMs for vision-language understanding and automated reporting.
- Implemented CI/CD pipelines using Docker and GitHub Actions for continuous training and deployment.
- Mentored an international AI team and collaborated with Japanese and English-speaking engineers to align product architecture with client goals.
- Proposed AI-driven product improvements contributing to business growth and operational efficiency.

AI Engineer (2020 - 2022)   Yangon, Myanamar

- Built real-time tracking and analytics systems with FastAPI, Flask, and MongoDB.
- Optimized deep learning models (YOLOX, NanoDet, OC-SORT) for multi-camera inference using TensorRT.
- Developed robust APIs and model-serving pipelines supporting distributed GPU workloads.

Projects

Multi-Camera Real-Time Baggage Tracking System

- Architected and led the development of a high-performance, real-time baggage tracking system across five synchronized surveillance cameras
- Integrated RT-DETR for high-accuracy baggage detection and OC-SORT for real-time multi-object tracking.
- Implemented Centroid-based ReID to ensure consistent baggage identity across non-overlapping camera views.
- Implemented concurrent multi-threading to achieve real-time processing across multiple video feeds.

AI-Powered Blueprint Symbol Detection System

- Built a scalable FastAPI backend with Celery and RabbitMQ for high-throughput, asynchronous distributed processing.
- Achieved 95%+ detection accuracy in low-data settings using a custom-trained YOLOX-m model on high-res TIFF images.
- Designed a synthetic data generation pipeline and integrated inference with a Qt C++ desktop application with MongoDB used as data storage.

Real-Time Person Re-Identification System

- Developed a real-time, full-stack tracking system to identify and follow individuals across multiple street-level CCTV cameras.
- Combined NanoDet for lightweight detection with OC-SORT for efficient multi-camera re-identification, achieving responsive performance.
- Engineered a full-stack web application using React, Flask, and MongoDB, enabling real-time data visualization from camera feeds and enhancing analytics processing speeds by 40%.

Multimodal Visual Understanding & Report Generation System

- Built a multimodal system using Streamlit and FastAPI to enable image interpretation and natural language generation for visual understanding.
- Automated structured Excel report generation with OpenPyXL, embedding input images and model-generated textual insights.
- Integrated Phi-3.5-Vision-Instruct and Phi-3.5-Mini-Instruct from HuggingFace for efficient local inference, combining VLM and LLM capabilities.

Certifications

- **AWS Machine Learning Specialty**
- **AWS Cloud Practitioner**
- **DeepLearning.AI Specialization in Deep Learning**

Languages

- **English - Business Level**
- **Japanese - N3**