NGUYEN VAN TU

AI Engineer Intern

0964511270 nvtu2305@gmail.com linkedin.com/in/nguyen-van-tu github.com/Tuprott991

Education

VNUHCM - University of Science **2022 - 2026**

Bachelor of Information of Technology — Major: Computer Vision *Ho Chi Minh City*

- **Courseworks**: OOP, DSA, Databases, Mathematics for AI, Probability and Statistics, AI Fundamentals, Machine Learning, Digital Image & Video Processing
- **Foreign language:** IELTS 6.0

Honors & Awards

- **Champions** of Web3 & AI Ideathon (2025) (over 450+ teams)
- **Finalist** in AI Challenge HCMC 2024
- **Champions** of Line Follower Robot competition HCMUS (F-RACE) 2024
- **Consolation Prize** in The National Youth Informatics Competition 2022

Publication

Tu et al. **An Interactive System For Visual Data Retrieval From Multimodal Input**. *The International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making (IUKM 2025).*

- Work under Dr.Dang Bui from Sep to Nov 2024 to develop a conversational and multimodal video event search event.
- Research and apply AI models like CLIP, Whisper, PaddleOCR, TransnetV2, and GPT-4o, allowing users to retrieve visual data using natural language, image, and voice.

Technical Skills

- **Languages**: Python, C++, JavaScript
- **Libraries/Frameworks**: FastAPI, LangChain, Pytorch, YOLO, Transformers, Scikit-learn, Pandas, Matplotlib, Milvus
- **Developer Tools**: Azure, Git, Docker

Projects

Educhain | 5 members | *github.com/Tuprott991/Educhain-AI* **Jan – Mar 2025**

- **Role**: Team lead / AI Engineer
- **Description:** Built a personalized learning platform integrating AI agents for **lightRAG-chatbot**, **study set generation**, **knowledge profiling**, and **graph file processing**.
- Utilized **FastAPI** for fast and efficient request handling backend. Leveraged **Langchain** to build agents with optimized retrieval and LLM-database interactions.
- Managed all data with **PostgreSQL**, leveraging **pgvector** for vector storage and **Apache AGE** for knowledge graphs
- Fine-tuned **Qwen2.5-7B** using the **LoRA** method and deployed it with **vLLM** for optimized inference.
- **Techs:** Python, FastAPI, Langchain, PostgreSQL, LoRA, vLLM, lightRAG.

Multimodal Video Retrieval | 5 members | *github.com/Tuprott991/AIthena-C* **Aug – Oct 2024**

- **Role:** Team lead / AI Engineer
- **Description:** Developed a **AI-Driven** system to search videos event based on natural language, scene, voice, OCR, and other metadata.
- Reduced search latency by **30%** through optimized keyframe extraction using **OpenCV** and **TransNetV2**.
- Applied **CLIP ViT L/14** and **BLIP-2** for embedding generation, enabling efficient vector search with **FAISS**.
- Enhanced retrieval performance with **multimodal inputs** (text, voice, prompts, objects); integrated **GPT-4o** for query refinement and visual question answering, and employed **Whisper** for accurate real-time speech-to-text conversion.
- **Techs:** Python, Flask, HuggingFace, Numpy, Transformer, OpenAI

Eventlens - Photo Album Event Recognition | *github.com/Tuprott991/EventLens* **March - April 2025**

- **Description:** Developing and training a transformer-based model for multi-label photo album event recognition.
- Implemented a novel architecture leveraging the **Vision Transformer** to extract features and apply an **attention mechanism** to images in an album for better context awareness.
- Achieved **97.86%** mean averange precision (**State-of-the-art**) for CUFED validate dataset
- **Techs:** Python, Pytorch, Transformers, Huggingface, Scikit-learn, torchvision

Vision Language Object Tracking | *github.com/Tuprott991/Object-tracking-Natural-Language* **Dec 2024**

- **Description:** Designed an object tracking pipeline that integrates natural language inputs for enhanced object identification and tracking across video frames. Leveraged vision-language models for understanding scene context and tracking targets based on user queries.
- Improved object tracking precision by 20% with a custom-trained **YOLOv5** model on a Vietnamese vehicle dataset.
- Implemented a hybrid approach combining candidates matching with **CLIP** for semantic query understanding and **DeepSORT** for robust multi-object tracking.
- **Techs:** Python, Ultralytics, YOLO, CLIP, DeepSORT, Googletrans