Thang-Long Nguyen-Ho

✉ contact@nhtlongcs.com | ⌂ github.com/nhtlongcs
in linkedin.com/in/nhtlongcs | G google scholar

EDUCATION

Ph.D. Student, Dublin City University (DCU) - Full Scholarship   2024 - Present
Research Topic: Information Retrieval
- Supervised by Prof. Cathal Gurrin and Prof. Graham Healy
- Conducting research in multi-media retrieval systems with a focus on representing temporal relationships and providing explainable answers through logical agent interpretations using Large Language Models (LLMs).

B.S Computer Science, University of Science, VNU-HCM - Honors Program   2018 - 2023
Major GPA: 3.89 / 4.00
Thesis - Semi-supervised Organ Segmentation - 4.00 / 4.00
- Supervised by Prof Minh-Triet Tran
- Proposes an improved organ segmentation model for volumetric CT images using semi-supervised methods and active learning strategies.

WORK EXPERIENCE

AI Engineer (Contractor), PiMA Academy   2025
- Led the quality assurance of a comprehensive mathematics dataset, delivering over 15,000 high-quality problems with solutions, in collaboration with the xAI team to enhance language model capabilities.
- Spearheaded the organization of annual mathematics summer camps focused on Deep Learning, Mathematics in Data Science, Bioinformatics, and Linear Programming for high school students.

Researcher, SELab, University of Science, VNU-HCM   2022 - 2024
- Conducted model architecture research in object recognition, tracking, and event retrieval from videos
- Implementing, training and finetuning vision models, focusing foundation models and large scale datasets.
- Leveraged deep learning frameworks like PyTorch and TensorFlow for model development and optimization, with a focus on embedding models (CLIP, SigLIP) and vector databases (FAISS, Milvus) for data retrieval.

AI Engineer Internship, MAPDAS   2021 - 2022
- Designed and implemented a custom PyTorch object detection model for identifying and classifying traffic signs within challenging 360-degree camera imagery.
- Led the curation and development of a Vietnamese traffic sign dataset, comprising over 5,000 annotated images across more than 20 distinct categories.

Research Assistant, Robotics & IoT Club, AI Lab, University of Science, VNU-HCM   2019 - 2020
Supervised by MSc Xuan-Nam Cao, Assoc. Prof Minh-Triet Tran on Robotics, Computer Vision
- Algorithm designing for autonomous car system control and navigation, develop and optimize deep model processing time.

SELECTED PROJECTS

Lifelog Retrieval System (WIP): An AI-powered system for human memory organization   2025
- Open source as part of my PhD research, focusing on delivering paper results into a real-world application
- Utilizes logical agent framework (Pydantic-AI) and Gemini for handling multiple types of temporal questions and leveraging powerful reasoning capabilities of LLama3.2 vision for multi-media understanding
- Designed with a FastAPI backend, the system is engineered to support multiple downstream tasks

StarListify: Developed a tool to curate and manage starred repositories on GitHub, enhancing knowledge organization for developers.   2024
- Implemented features to categorize and tag repositories, making it easier for users to navigate and access their favorite projects.
- Reverse engineered the GitHub API to discover unofficial endpoints, allowing language models to analyze user behavior and repository relationships.

Search Services: Developed a comprehensive solution for running searching and indexing services using Docker, enhancing scalability and performance.   2023
- Created and configured Docker images for search services, utilizing technologies like Elastic Search and Milvus for efficient image and text search capabilities.
- Developed the PySearch SDK, a Python interface to facilitate communication with the services, improving usability and integration.

- Achieved seamless deployment and management of services through Docker Compose, optimizing resource usage and operational efficiency.
- Implemented a continuous integration and continuous deployment (CI/CD) pipeline using GitHub Actions, ensuring automated functionality testing, building, and deployment of services.

Vietnamese OCR Toolbox: A comprehensive toolbox that provides a set of tools for all essential steps of an OCR pipeline, from image pre-processing to text extraction and recognition, effective in recognizing Vietnamese text in a variety of documents. 2021
- Include document extraction and normalization algorithms from wild images.
- Provide a training pipeline for Vietnamese text recognition, and multiple text region detection and Vietnamese OCR algorithms from a variety of libraries.

SSDF - Simulation Self Driving Framework: Project aiming at solving autonomous driving problems in a virtual simulation. 2020
- It provides a collection of Pytorch algorithms for perception, prediction, and control, as well as tools for data processing and training.

SELECTED PUBLICATIONS

Conference - Workshop Papers
- **Thang-Long Nguyen-Ho**, Minh-Khoi Pham, Tien-Phat Nguyen, Hai-Dang Nguyen, Minh N. Do, Tam V. Nguyen, Minh-Triet Tran, 2022. **Text Query based Traffic Video Event Retrieval with Global-Local Fusion Embedding**. *The 2022 Ai City Challenge - CVPR* - **[Code]**

SELECTED AWARDS

First Prize in Huawei Cloud Optimization Competition 2024
- Solved a complex cloud resource management problem involving the efficient management of servers across data centers, aiming to maximize the profit.
- Competed with international teams and earned the top prize €6,000 by designing an explainable, real-world solution tailored to industrial challenges in cloud optimization.

First Prize in UCC AI Quest 2024
- Developed an model for segmentation vegetation patches in natural landscapes using high-resolution aerial images from drones.
- Awarded €5,000 for best team solution, competing against researchers from diverse backgrounds.

First Prize on Ho Chi Minh AI Challenge 2020
- Designed a vehicle counting algorithm that follows pre-defined movements from multiple camera scenes in a Vietnamese dataset. Contestants included researchers and engineers at all levels in Vietnam.
- Awarded €4,000 for best team performance.

First Prize on Ho Chi Minh AI Hackathon 2020
- Awarded €1,000 for best team performance in detecting anomaly events from cameras in Vietnam.

First Prize International Science and Engineering Fair, High School, Software Engineering field 2017
- Guaranteed admission to national universities.
- Implemented an interactive application with hand gestures to visualize virtual anatomy and simulate interactions with physical materials.
- Intel Excellence Award in Computer Science.

SKILLS

Languages: English (fluent), Vietnamese (native)
Programming languages: Python, C++, JavaScript
Tools: Docker, Git, CMake, Slurm
Deep learning frameworks: PyTorch, LibTorch, TensorRT, Huggingface, TensorFlow
Data Technologies: SQL, NoSQL, MongoDB, Milvus, Elastic Search, Pydantic
Machine Learning and AI: NLP, Large Language Models (LLMs), Embedding Models, Model Training, Inference Models, Model Deployment, High-performance computing (HPC)
Agentic AI Concepts and Frameworks: Large Language Models (LLMs) like GPT, Function Calling, RAG, Chain of Thoughts, Pydantic AI, LlamaIndex, Langchain, Agno
Development Practices: Focus on data structures and design patterns, CI/CD