Nhat M. Hoang

Singapore, 644276 | (+65) 85877324 | nhat005@e.ntu.edu.sg | nhathoang2002.github.io

EDUCATION

Nanyang Technological University (NTU) Aug 2020 – May 2024
Bachelor of Engineering in Computer Science Singapore

- Elective Specializations: Artificial Intelligence; High Performance Computing
- Recipient of MOE Tuition Grant with a **3-year bond of working in Singapore.**

RESEARCH PUBLICATIONS

- *ToXCL: A Unified Framework for Toxic Speech Detection and Explanation*
  Nhat M. Hoang*, Xuan Long Do*, Duc Anh Do, Duc Anh Vu, Luu Anh Tuan
  NAACL, 2024 (Main Technical Track)
- *MotionMix: Weakly-Supervised Diffusion for Controllable Motion Generation*
  Nhat M. Hoang, Gong Kehong, Chuan Guo, Michael Bi Mi
  AAAI, 2024 (Main Technical Track)
- *ChatGPT as a Math Questioner? Evaluating ChatGPT on Generating Pre-university Math Questions.*
  Phuoc Pham Van Long*, Duc Anh Vu*, **Nhat M. Hoang***, Xuan Long Do*, Anh Tuan Luu
  ACM/SIGAPP SAC, 2024 (Main Technical Track)
- *Data Augmentation using Corner CutMix and an Auxiliary Self-supervised Loss.*
  Fen Fang, **Nhat M. Hoang**, Qianli Xu, Joo-Hwee Lim
  ICIP, 2023 (Main Technical Track)

SKILLS

Languages: Python, Jupyter Notebook, C/C++, SQL, Java, HTML, CSS, JavaScript, MATLAB.

Technologies: PyTorch, Transformers, Pandas, NumPy, Matplotlib, Scikit-Learn, OpenCV, TensorFlow, Keras, GitHub, Shell Scripts, AWS, PostgreSQL, MongoDB, Docker, NodeJS, ReactJS, Blender.

PROFESSIONAL EXPERIENCE

Huawei Singapore Research Center Mar 2023 – Nov 2023
Algorithm Engineer Intern Singapore

- Researched on utilizing Multi-modal Large Language Model for interactive 3D chatbot.
- Overcame the scarcity of high-quality data in controllable 3D-mesh full-body human motion generation by designing an innovative weakly-supervised diffusion model to leverage sources of low-quality data.
- **Attained up to a 39% improvement** in generating diverse and realistic human motion across tasks like text-to-motion, action-to-motion, and music-to-dance through training diffusion models with Python and PyTorch.

NTU NAIL Lab Oct 2022 – Present
Student Research Assistant Singapore

- Collaborated with a group of students to assess the performance of 5+ pre-trained language models (e.g., GPT-2, T5) and prompting Large Language Model (e.g, ChatGPT, LLaMa) in generating math word problems on 4 benchmarks.
- Led the development of a pioneering end-to-end framework for implicit hate speech detection and explanation, achieving state-of-the-art results in both tasks by fine-tuning pretrained language models using Python and PyTorch.

I2R Department, ASTAR Jun 2022 – Dec 2022
Computer Vision Research Intern Singapore

- Designed an innovative data augmentation approach for 2D image processing and incorporated a novel auxiliary loss to boost the generalization performance of SimCLR, a self-supervised framework.
- **Increased up to 8% in top-1 classification accuracy** across multiple benchmarks, including STL10, CIFAR100, and Food101, through model training with Python and PyTorch. Research work is accepted at **ICIP 2023.**

Eureka Robotics Jan 2022 – Apr 2022
Computer Vision Engineer Intern Singapore

- Deployed a stereo-matching model on AWS utilizing Robotics Operating Systems (ROS) Docker image, enabling the generation of disparity maps from stereo image pairs and, consequently, the creation of 3D point clouds.
- **Synthesized 1000+ images** using only 30 in-house collected image pairs, accomplished through the implementation of a pipeline in Blender.
- **Achieved an average end-point error below 3%** by fine-tuning state-of-the-art models with Python and PyTorch on the 1000+ synthesized images.

Ubisoft Jul 2021 – Oct 2021
Data Scientist Intern Singapore

- **Optimized the ranking performance from 55% to 89% AUC** through a collaborative effort, leveraged data analysis on three large-scale datasets to identify pivotal factors influencing recommendation performance.
- Implemented and evaluated state-of-the-art machine learning algorithms, analyzed trade-offs between personalized and popularized recommendations to ultimately select the most fitting model to deploy.

ACADEMIC PROJECTS

Visual Kinship Recognition | Coursework Project Mar 2023 – May 2023

- Analyzed a dataset comprising 20K+ facial images from 1000+ disjoint families to conduct feature engineering for the challenge of kinship recognition.
- **Achieved a 2nd place ranking** out of 500+ teams, resulting in an A grade, by training 20+ models in Python and Keras, employing techniques such as feature fusion and weighted ensemble methods.

Vietnamese Open-domain Question Answering | ZaloAI Challenge 2022 Nov 2022 – Dec 2022

- **Reduced article retrieval time from 3 seconds to 1 second per question** while maintaining performance by utilizing Python and SQLite library to efficiently slice and store the crawled Wiki articles in a local database.
- **Ranked among the top 10%** in the competition by engineering an industry-level end-to-end pipeline by ensembling multiple models and processing techniques.

Automatic Trash Classification | MLDA DLW 2021 Hackathon Oct 2021 – Oct 2021

- Led the development of an automatic trash classification system, employing a monocular camera to detect 6 classes of recyclable trash and seamlessly opening the corresponding trash bin.
- Fine-tuned ResNet50 model for precise classification of six recyclable classes leveraging Python and Keras. Contributed to integrating the model with a 6-LED system using C++ and Arduino for the demo submission.
- **Won the "Most Socially Impactful Hack"** out of 100+ teams in the hackathon.

AWARDS & ACHIEVEMENTS

Kaggle - Benetech Making Graphs Accessible | *Top: 22% out of 608 teams* Jun 2023
2022 ICPC Asia-Manila Programming Contest | *Rank: 18th out of 28 teams* Dec 2022
ZaloAI Challenge 2022 - E2E Question Answering | *Top: 10% out of 110 teams* Dec 2022
Deep Learning Week Hackathon 2021 | *Prize: Most Socially Impactful Hack out of 153 teams* Oct 2021
2nd Kibo Robot Programming Challenge | *Top: 6th World Finalist* Oct 2021
Shopee Code League 2021 - Data Science Challenge | *Top: 25% out of 1034 teams* Mar 2021
American Mathematics Contest 8 - AMC 8 | *Prize: Second Place* Aug 2015

CO-CURRICULAR ACTIVITIES

IEEE ISMAR 2022 Conference Singapore
Student Volunteer Oct 2022 – Oct 2022

- Assisted 50+ online participants with a better experience with the platform GatherTown.

Vietnamese Youth Alliance in Singapore (VNYA) Singapore
Publicity Department Member Sep 2021 – Sep 2023

- Promoted 100K+ reach on Facebook page through content creation, poster design of 100+ year-round event posts.
- Co-organized two internal events for 60+ members and three external events with 100+ participants each.