Hao Phung

hao-pt.github.io | tienhaophung@gmail.com | hao-pt | Google Scholar

Research interests

My primary research interests lie in the field of Computer Vision, with a specific focus on deep generative models. Currently, I am actively engaged in improving the efficiency and controllability of diffusion models, particularly in their application to conditional image generation.

Education

- **Vietnam National University Ho Chi Minh City - University of Science** — Viet Nam
  Bachelor of Computer Science; GPA: 8.40/10 *(in-major GPA:* 9.05/10*); Rank:* 14/320 — *Aug 2016 – Nov 2020*
  - Coursework: Computer Vision, Machine Learning, Artificial Intelligence, Data structures & Algorithms.

  Honors thesis: Human action monitoring based on Visual question answering — *Aug 2019 - Aug 2020*
  Advisor: Assoc. Prof. Ngoc Quoc Ly; Grade: 10/10
  - Utilized human-prior knowledge formed as Task Ontology to instruct the system what visual tasks should be performed to produce a suitable answer for an input query under Surveillance context.

Publications

(*) denotes equal contribution

- Quan Dao*, Hao Phung*, Binh Nguyen, Anh Tran, **Flow Matching in Latent Space**, *arXiv preprint*, 2023. [paper, code, page]
- Thanh Van Le*, Hao Phung*, Thuan Hoang Nguyen*, Quan Dao*, Ngoc Tran, Anh Tran, **Anti-DreamBooth: Protecting users from personalized text-to-image synthesis**, in *International Conference on Computer Vision (ICCV)*, 2023. [paper, code, page]
- Hao Phung*, Quan Dao*, and Anh Tran, **Wavelet Diffusion Models are fast and scalable Image Generators**, in *Computer Vision and Pattern Recognition (CVPR)*, 2023. [paper, code]
- H. Vo*, T.H. Phung*, and N. Ly, **VQASTO: Visual Question Answering System for Action Surveillance based on Task Ontology**, in *NAFOSTED Conference on Information and Computer Science (NICS)*, 2020. [paper]

Experience

- **VinAI Research** — Hanoi, Vietnam
  AI Research Resident - Advised by Dr. Anh Tran — *Aug 2021 - now*
  - Project: Flow Matching in Latent Space
    - Introduce a latent flow matching framework that targets high-resolution image synthesis and various types of conditional image synthesis.
  - Project: Anti-DreamBooth: Protecting Users from Personalized Text-to-Image Synthesis
    - Introduce perturbation learning algorithms for enhanced user protection against malicious risks in personalized text-to-image synthesis.
  - Project: Wavelet Diffusion Models are fast and scalable Image Generators
    - Propose a wavelet-based diffusion scheme that accelerates image generation by leveraging low- and high-frequency components of wavelet subbands at the image and feature levels.
    - Achieve near real-time performance, effectively narrowing the speed gap with GAN counterparts.

  Applied Rotation Program - Led by Mr. Tin Trung Duong — *Jul 2022 - Oct 2022*
  - Present a pipeline for Object Search using Open Vocabulary Object Detection that enables the retrieval of similar outputs based on a query object.

AI Engineering - Advised by Dr. Toan Duc Bui and Dr. Rang Nguyen *Dec 2020 - Aug 2021*

- Investigated SOTA semi-supervised learning for Image Classification and Monocular 3D Object Detection.
- Validated and benchmarked AI models for Autopilot projects (e.g. Camera Degradation, Lane detection).

- **Skeleton-Based Abnormal Behavior Recognition** Ho Chi Minh, Vietnam
  Research Collaborator - Led by Assoc. Prof. Ngoc Quoc Ly (cooperated with SNA Global) *Sep 2019 - Mar 2020*
  - Developed a real-time anomaly action recognition system by enhancing time efficiency for pose tracking and transforming skeleton sequence to image as new spatio-temporal feature for action recognition.
  - Improved accuracy of action recognition model by leveraging EfficientNet models and built up a minimal website using Flask framework for demonstration.

- **KMS Technology** Ho Chi Minh, Vietnam
  AI Engineer Intern - Advised by Mr. Hoa Trong Vu *Aug 2019 - Nov 2019*
  - Created new dataset (nearly 3000 images) by collecting and refining images manually for Image matching problem. Also, utilized Image Hashing algorithm for filtering out similar samples.
  - Increased accuracy by 2.5% (at 96.5%) through fine-tuning state-of-the-art ImageNet models and adding new augmentation methods on the new dataset.

Open-source Projects

- **Automatic License Plate Recognition (ALPR)** *Apr 2019 - Jun 2019*
  - Redesigned a handcraft algorithm for ALPR by adding FloodFill algorithm to extract plate more precisely and post-processing character image in plate to increase the accuracy of plate number recognition.
  - Extended ALPR for motorbike plate recognition by utilizing Haar-Cascade Classifier to detect the plate.
  - **Tech stack**: Python, OpenCV, Haar Cascade Classifier, SVM, Tesseract-OCR, Git.
- **Face Recognition** *May 2019 - Jun 2019*
  - Built a model for face recognition through utilizing cutting-edge face detection methods (e.g. SSD, Multitask-CNN) and optimizing the SVM model for identity recognition on our dataset.
  - Adopted FaceNet for feature extraction as input to SVM model.
  - **Tech stack**: Python, OpenCV, SSD, FaceNet, SVM, Facial landmarks, Git.

Professional activities

- Reviewer: TPAMI (2023)

Honours and Awards

- Outstanding thesis award *2020*
- Top 5 IT students in academic year *2018 - 2019*

Programming Skills

- **Languages**: Python, C/C++, HTML/CSS, SQL.
- **Technologies**: PyTorch, Tensorflow, OpenCV, Scikit-learn, Git, LaTeX, Docker, Linux.

Languages

- Vietnamese: Native
- English: IELTS Academic 7.0
- Cantonese: Beginner

Extracurricular Activities

- **AI Day 2022** Hanoi, Vietnam
  Panel speaker *Aug 2022*