Nguyen (William) Nguyen

Palo Alto, CA, 94303
Email: nguyennm1024@gmail.com | LinkedIn | GitHub | Google Scholar | Personal Website : (+1) 585-540-6264

EDUCATION

University of Rochester, Rochester, NY 08/2022 – 05/2024
M.S. in Computer Science

University of Engineering and Technology, Hanoi, Vietnam 08/2016 – 08/2020
B.S. in Computer Science

RESEARCH INTERESTS

Vision-Language: Visual captioning; Multimodal-LLMs; Scene text understanding; Language-guided visual generation.

Machine Learning: Representation learning; Unsupervised learning; Self-supervised learning.

WORK EXPERIENCE

Senior Applied Scientist, Aitomatic, Inc, Palo Alto, CA 08/2024 – now
Reported to: Dr. Christopher Nguyen (CEO)

- Domain-specific foundation model: Led the development of SemiKong and Llamarine, the first open-sourced LLMs for semiconductors and maritime domains, which were used by TSMC, Tokyo Electron, and Furuno.
- TRAIN product: Led the technical development of DXA factory: A system allows users to create domain-specific expert agents in a few clicks. Proposed and implemented novel techniques of automated (pretraining and finetuning) data curation for LLMs training and expert knowledge synthesis for agent training, **saving 5x the total training cycle time**, helping the company to serve 20% more customers than before.
- Agents: Developed ProSEA, a hierarchical structured agent solves the problems by learning from past failures.
- Research and Development: Authored 2 research papers and 3 US patent applications.

Research Assistant, University of Rochester, Rochester, NY 06/2022 – 06/2024
Advisor: Professor Chenliang Xu

- Object state captioning: Propose a new task to describe object states in detail. Built a model using multimodal-LLM that can perform QA, conversation, and reasoning.
  Model **achieved 90% compared to GPT4V** on both metrics and human evaluation.
- Scene text spotting: Incorporate language priors to make the Scene text spotting system more robust, significantly surpassing SOTA from 2-4% in every standard benchmark.
- Instructional video understanding: Understand human action and building task guidance algorithms from egocentric and instructional videos. [Demo video].

AI Research Resident, VinAI, Hanoi, Vietnam 12/2019 – 06/2022
Advisor: Professor Nguyen Minh Hoai

- Scene text recognition: Incorporate knowledge from a dictionary into both the training and inference stage, surpassing the SOTA by 3-5%. Introduce a novel Vietnamese scene text understanding dataset.
- Scene text spotting for street sign: Develop a framework for text traffic sign recognition with improved inference speed, proposing an annotation method saves the company 50% annotation cost and maintains data quality.
- Face recognition: Develop unified models for recognizing normal faces, faces with masks, and extreme pose faces. Improve masked face recognition by 18% by generating synthetic masked face images from normal faces. Build a lightweight model for running on-edge devices by using knowledge distillation.

Research Assistant, Vietnam National University, Hanoi, Vietnam 06/2018 – 03/2020
Advisor: Professor Xiem Hoang Van

- Machine learning for video coding: Define handcraft feature and use machine learning to speed up quad-tree partitioning and enhance decoded frame quality by classifying whether blocks need to be splitted.
- Co-advise junior students: Support junior students to develop a project: ID card information extraction.

AI Engineer Intern, Teko, Hanoi, Vietnam 04/2019 – 11/2019

- Product Clustering: Represent e-commerce product data by attributes and product descriptions, building an automated pipeline from data collection, feature selection, dimensionality reduction with VAE, and clustering. Build a visualization dashboard for product analysis.

TECHNICAL SKILLS

Programming Languages Python, C/C++
Deep Learning Framework PyTorch, TensorFlow, Scikit learn, MLlib
Others OpenCV, PySpark, Pandas, Matplotlib, Numpy, Flask
Tools Bokeh, Git, Docker

PROFESSIONAL SERVICES

Reviewer: WACV 2022, CVPR 2023, CVPR 2024, ACM MM 2024, AAAI 2024, CVPR 2025
Invited Speaker: VinAI Research Workshop 2021.
Organizer: Vietnamese Scene Text Recognition Challenge 2021.

PUBLICATIONS

1. **William Nguyen**, An Phan, Konobu Kimura, Hitoshi Maeno, Mika Tanaka, Quynh Le, William Poucher, Christopher Nguyen, "*Llamarine: Open-source Maritime Industry-specific Large Language Model*", preprint, 2024.

2. Christopher Nguyen, **William Nguyen**, Atsushi Suzuki, Daisuke Oku, Hong An Phan, Sang Dinh, Zooey Nguyen, Anh Ha, Shruti Raghavan, Huy Vo, Thang Nguyen, Lan Nguyen, Yoshikuni Hirayama, "*SemiKong: Curating, Training, and Evaluating A Semiconductor Industry-Specific Large Language Model*", OSAI4MU, AAAI, 2025.

3. Vinh Luong, Sang Dinh, Shruti Raghavan, **William Nguyen**, Zooey Nguyen, Quynh Le, Hung Vo, Kentaro Maegaito, Loc Nguyen, Thao Nguyen, Anh Hai Ha, Christopher Nguyen, "*DANA: Domain-Aware Neurosymbolic Agents for Consistency and Accuracy*", preprint, 2024.

4. Jing Bi, Yunlong Tang, Luchuan Song, Ali Vosoughi, **Nguyen Nguyen**, Chenliang Xu, "*EAGLE: Egocentric AGgregated Language-video Engine*", ACM Multimedia (ACMMM), 2024.

5. **Nguyen Nguyen**, Jing Bi, Ali Vosoughi, Yapeng Tian, Pooyan Fazili, Chenliang Xu, "*OSCaR: Object States Captioning and State Changes Representation*", NAACL, 2024.

6. Jing Bi*, **Nguyen Nguyen***, Ali Vosoughi*, Chenliang Xu (* equal contribution), "*MISAR: A Multimodal Instructional System with Augmented Reality*", AV4D, International Conference on Computer Vision (ICCV), 2023.

7. **Nguyen Nguyen**, Yapeng Tian, Chenliang Xu, "*Efficiently Leveraging Linguistics Knowledge for Scene Text Spotting*", under reviewed, 2023.

8. **Nguyen Nguyen**, Thu Nguyen, Vinh Tran, Triet Tran, Thanh Duc Ngo, Thien Nguyen, Minh Hoai, "*Dictionary-guided Scene Text Recognition*", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2021.

PATENTS

1. Christopher Nguyen, **Manh-Nguyen Nguyen**, Hong An Phan, Zooey Nhu-Quynh Nguyen, The-Vinh Luong, Elise Nhu-Y Nguyen, Thomas Rasmussen, Anh Hai Ha, Phi-Hung Vo, Xuan-Sang Dinh, Huy-Thuan Bui, Anh-Quoc Dang, Timothy Michael Gerard Rozario, "*Delivering Domain-Expert Agents and Models Using Synthetic Knowledge*", US Patent App. 63/726,322, 2024.

2. Christopher Nguyen, **Manh-Nguyen Nguyen**, Hong An Phan, Zooey Nhu-Quynh Nguyen, The-Vinh Luong, Elise Nhu-Y Nguyen, Thomas Rasmussen, Anh Hai Ha, Phi-Hung Vo, Xuan-Sang Dinh, Huy-Thuan Bui, Anh-Quoc Dang, "*Delivering Domain-Expert Agents for Improving Problem-Solving*", US Patent App. 63/721,419, 2024.

3. Christopher Nguyen, The Vinh Luong, Xuan Sang Dinh, Zooey Nhu-Quynh Nguyen, Shruti Raghavan, **Manh Nguyen Nguyen**, Quynh Thi-Tham Le, Phi Hung Vo, Tan Loc Nguyen, Anh Hai Ha, Phuong Thao Nguyen, "*Domain-Aware Neurosymbolic Agents For Improving Problem-Solving Accuracy And Consistency*", US Patent App. 63/696,337, 2024.