Nguyen (Will) Nguyen

Rochester, NY, 14611
Email: nguyennm1024@gmail.com | LinkedIn | GitHub | Google Scholar | Personal Website : (+1) 585-540-6264

EDUCATION

University of Rochester, Rochester, NY — 08/2022 – now
Ph.D. in Computer Science

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

PUBLICATIONS

1. **Nguyen Nguyen**, Jing Bi, Ali Vosouughi, Yapeng Tian, Pooyan Fazili, Chenliang Xu, "*OSCaR: Object States Captioning and State Changes Representation*", under reviewed, NAACL, 2024.
2. 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.
3. **Nguyen Nguyen**, Yapeng Tian, Chenliang Xu, "*Efficiently Leveraging Linguistics Knowledge for Scene Text Spotting*", under reviewed, 2023.
4. **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.

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

Research Assistant, University of Rochester, Rochester, NY — 08/2022 – now
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.
Invited Speaker: VinAI Research Workshop 2021.
Organizer: Vietnamese Scene Text Recognition Challenge 2021.