NGUYEN MANH NGUYEN

nguyennm1024.github.io nguyennm1024@gmail.com
Rochester, New York, United State (+1)-585-540-6264

RESEARCH INTERESTS

My research interest lies in making machine learning models that can understand vision and language information, reasoning, and extracting knowledge from images and videos. Additionally, I am also fond of few-shot learning. Specifically, my works focus on how to make the vision model more robust by leveraging language knowledge.

EDUCATION

University of Rochester, Rochester, New York *Aug 2022 - now*
Ph.D. in Computer Science

University of Engineering and Technology - Vietnam National University *Aug 2016 - Aug 2020*
Bachelor of Information Technology

RESEARCH EXPERIENCE

University of Rochester *Aug 2022 - now*
Ph.D. Student
- Supervisor: Prof. Chenliang Xu (google scholar)
1. Incorporating language knowledge to make the Scene text recognition system more robust.
2. Understanding human action and building a task guidance algorithm from egocentric and instructional videos.

VinAI Research *Dec 2019 - Dec 2021*
AI Research Resident
- Supervisor: Prof. Minh Hoai Nguyen (google scholar)
1. Proposing a new approach to improve h-mean score of state-of-the-art scene text spotting backbones using language prior. This method can be used in training and testing, improving many different backbones. This work has been accepted in **CVPR 2021**.
2. Introducing a novel Vietnamese scene text dataset (VinText) - the largest scene text dataset in Vietnam with 2000 fully annotated images and about 56.000 text instances.

VinAI Research *May 2021 - Sep 2021*
Applied Rotation Program
1. Collecting a text street signs dataset in videos. Proposing a reasonably priced annotation method but still maintaining data quality.
2. Developing a novel framework for text traffic signs recognition and improving inference speed.

Artificial Intelligence and Multimedia Signal Processing Lab, UET-VNU *Jun 2018 - Mar 2020*
Research Assistant
- Supervisor: Prof. Hoang Van Xiem (google scholar)
1. Using machine learning to speed up quad-tree partitioning and enhance decoded frame quality.
2. Co-supervising junior students and helping them to develop their project: ID card information extraction.

PUBLICATIONS

Can Linguistic Knowledge From a Large Text Corpus Help Scene Text Spotting? *2023*
***Nguyen Nguyen**, Yapeng Tian, Chenliang Xu*
ICCV 2023 Under Review.

Dictionary-guided Scene Text Recognition *2021*
***Nguyen Nguyen**, Thu Nguyen, Vinh Tran, Minh Triet-Tran, Thanh Duc Ngo, Thien Huu Nguyen, Minh Hoai*
Conference on Computer Vision and Pattern Recognition, 2021.

INDUSTRIAL EXPERIENCE

VinAI Research *Jan 2022 - June 2022*

AI Research Engineer

1. Doing research, developing, and optimizing AI algorithms for face recognition problems. Developing unified models for recognizing normal faces, faces with masks, and extreme pose faces.
2. Building a lightweight model for running in edge devices by knowledge distillation

Teko Vietnam *Apr. 2019 - Nov. 2019*

AI Engineer Intern

1. Exploring e-commerce problems and using AI for several tasks: Product clustering, Customer segmentation, and Automatic keywords generation.
2. Working with e-commerce big data system, writing API service, building web-based tool to analyze products, and setup CI/CD

PINGCOM *Sep. 2018 - Jan. 2019*

AI Engineer Intern

1. Setting up Cassandra and PySpark with docker to manage Facebook data of over 100.000 users.
2. Developing name2gender models to predict gender from user names, then write an API using Flask

PROJECTS

Progress Estimation

- Estimating percentage of task progress using weakly supervised learning.
- Proposing few-shot learning approaches using language guidance.
- Tools: Pytorch, Torchvision, OpenCV.

Masked Face Recognition

- Applying the 3D face model technique to generate faces with masks from a single RGB image.
- Developing lightweight model with knowledge distillation.
- Verifying model robustness with matching score distributions comparison.
- Building a visualization tool for matching verification. This tool has been used in other teams for debugging.
- Tools: Pytorch, OpenCV.

Text traffic signs recognition

- Collecting a video text street sign dataset in Vietnam, proposing methods to annotate data more efficiently, and validating data quality. Developing deep learning models to recognize text in street signs and incorporating it into traffic sign recognition system for autopilot.
- Tools: Pytorch, OpenCV.

Scene text spotting

- Proposing a novel approach to incorporate language knowledge from the dictionary into training and testing stages, helping many models improve themselves without increasing complexity in the inference phase. Building an annotation tool for data collection and introducing a novel Vietnamese scene text dataset (VinText) and publicizing it to the community.
- Tools: Pytorch, OpenCV, javascript

Product Clustering

- Representing Phong Vu product data by attributes and product descriptions, building an automated pipeline from data collection, feature selection, dimensionality reduction with VAE, and clustering. Building a visualization dashboard for product analysis.
- Tools: PySpark, Bokeh, Scikit learn, Keras, Tensorflow, Pandas, Matplotlib, SciPy.

TECHNICAL SKILLS

Programming:	Python, C/C++
Software & Tools:	**Machine learning:** Pytorch, Keras, Scikit learn, MLlib
	Computer vision: OpenCV
	Data engineering: PySpark, Pandas, Matplotlib, Numpy, Flask
	Tools: Bokeh, Git, Docker

ACTIVITIES

Competition jury member *2021*
Ho Chi Minh city AI Challenge 2021: Vietnamese Scene Text Recognition

Invited speaker *2021*
Public computer vision workshop organized by VinAI Reseasrch

Teaching Assistant *Jun 2018 - April 2019*
University of Engineering and Technology - Vietnam National University
Teaching assistant in several computer vision and machine learning courses for Samsung Display Vietnam's staff