

VINH D. VAN



PERSONAL INFORMATION

email vinh.van@tuni.fi
phone +358 465 741 540
github [vinhvd749](#)

WORK EXPERIENCE

Research Assistant

May 2026 TAMPERE UNIVERSITY
Dec 2026

COMPUTER VISION GROUP

Supervisors: Prof. Esa Rahtu and Prof. Joni Kämäräinen

- LiDAR-free 3D forest mapping: a 1 km walk with a consumer 360° camera becomes a photorealistic, georeferenced reconstruction, with stems within 4–5 cm and ground within 11 cm of LiDAR.
- Reconstruction time cut from 13 hours to 2 hours at nearly the same quality (−0.5 dB).
- Streaming reconstruction that keeps up with the capture on one shared GPU and reaches the offline quality.
- Sensor calibration and timing for a LiDAR + camera + IMU rig.
- Benchmarked 3D foundation models (VGGT, Depth Anything 3); trunk and terrain labels for robot navigation in rough terrain.

Teaching Assistant

Oct 2025 TAMPERE UNIVERSITY
Dec 2025

DATA.ML.200: DEEP LEARNING

Course lecturer: Prof. Tuomas Virtanen

- Grading and providing feedback for students' exercises.
- Answering students' questions related to the course's content.
- Improving the reliability of the auto-grader code-base.

Data Scientist

Sep 2022 MoMo (M_SERVICE)
Jul 2024

GEOMAP (MOMO INTERACTIVE MAP)

- The goal of this project is to replicate all Google Maps features for MoMo clients and employees, resulting in substantial (several thousand dollars) monthly savings for the company.
- I crawled large numbers of locations (buildings, areas ...) from Google Map API, Grab, and Foody.vn. Then, these data sources were merged with in-house MoMo merchant data through location identity resolution.
- The resulting service provides in-house geocoding, reverse-geocoding, address auto-complete, interactive map, and geospatial data analytics.

Data Scientist

Feb 2020 FREELANCE WORK — VIETBIBLE
Jun 2021

ENGLISH-VIETNAMESE BOOK TRANSLATION SYSTEM

- I built a web application for translating English Word documents to Vietnamese. On the IWSLT2015 English-Vietnamese dataset, the system BLEU score was ≈ 38.7 , which was the state-of-the-art at the time (although this system was built for a specific domain requested by the customers). Here is [a translation sample](#).

- I crawled a large number of Vietnamese and English books online and matched the translation using LASER with clustering and sequence alignment algorithms. For Vietnamese data for which the English translation can not be found, I used the free Google Translate API.
- The resulted corpus contains ≈ 17 mil sentence pairs. I also extracted sentence grams from paragraphs (from 2-5 consecutive sentences), which resulted in more than ≈ 68 mil training samples. The created dataset was used to train a Transformer.

Associate Data
Scientist

Oct 2017
Feb 2020

VNG — INFORMATION SECURITY OFFICE (ISO)

ZALOPAY FRAUD DETECTION

- When I joined the project, a very small number of fraud cases had been reported, which includes promotion farming and credit card fraud. By analyzing the graph of shared accounts and devices, and money flow between ZaloPay accounts, I detected $\approx 5541\%$ more than the number of reported fraud cases, these cases were then investigated and confirmed by ZaloPay. This analysis uncovered many fraud schemes that had not been known before.
- As the graph methods had proven their versatility, I was requested by ZaloPay to build a Tableau dashboard for fraud monitoring. Every 30 minutes, fraud reports would be sent to the ZaloPay Risk Management team, and the dashboard would be updated.
- In total, these methods had helped detect a significant amount of money lost to fraud, especially promotion farming, which could consist of up to $\approx 70\%$ awarded money for some promotion campaigns. This also helped ZaloPay to make quick adjective decisions for their later promotion campaigns.

Data Scientist
Intern

Sep 2016
Mar 2017

EYEQ TECH

OPEN-SET FACE RECOGNITION SYSTEM

- This system was crucial for funding [EyeQ Tech](#). I was working under Dr. Tung Le (who would be EyeQ Tech's founder) for ShareCarForAds at the time.
- Given a sequence of face tracking images, I built a system to recognize if the person has been registered or not. If registered, return his/her ID; otherwise, register this person with a new ID.
- I also tested and optimized the system's response time using [pprofile](#) as it has a significant effect on the system's UX.

PUBLICATIONS

SoICT 2017

Dec 2017

Combining Convolution And Recursive Neural Networks For Sentiment Analysis

[acm.org \(publishing site\)](#) · [researchgate.net \(PDF\)](#) · [github \(code\)](#)

Author: **Vinh D. Van**, Thien Thai, Quoc-Minh Nghiem

EDUCATION

M.Sc. in
Statistical
Data Analytics

2024-present

TAMPERE UNIVERSITY, TAMPERE

GPA: 4.95/5 · 100% scholarship · [Transcript](#)

B.Sc. in
Computer Science

2013-2017

UNIVERSITY OF SCIENCE, HO CHI MINH CITY

GPA: 8.0/10 · Program: [Advanced Program in Computer Science](#)

Languages

ENGLISH: [IELTS 7.5](#) · VIETNAMESE: mother tongue