NGUYEN QUANG VINH, M.Sc.

✉ vinhbn28@gmail.com   🐦 @VinhNguyen8220   in VinhNguyen

🌐 mypage   ⌨ vinhhust2806

AI Researcher

Aspiring AI researcher with a foundation in computer vision (CV) and natural language processing (NLP), backed by publications in prestigious AI conferences, including A-rank venues such as BMVC and ACCV. My research expertise encompasses medical image analysis, mental health detection, face generation, and media search, where I have developed novel deep learning models and conducted impactful experiments. Beyond academia, I have extensive hands-on experience in diverse AI projects, including chatbot development, depression detection, action localization, tracked-vehicle retrieval, and face anti-spoofing. Driven to advance cutting-edge AI research, I aim to apply my interdisciplinary knowledge to solve complex challenges and push the boundaries of innovation in artificial intelligence.

Employment History

2023 – 2025 🔖 **AI Research** Pattern Recognition Laboratory, Department of Artificial Intelligence Convergence, College of AI, Chonnam National University.
Pattern Recognition Laboratory.

2022 – 2023 🔖 **Leader** AI VIET NAM.
https://www.facebook.com/aivietnam.edu.vn.

2021 – 2022 🔖 **AI Research** Image Processing and Signal Analysis Laboratory, Department of Automation, School of Electrical and Electronic Engineering, Hanoi University of Science and Technology.
Image Processing and Signal Analysis Lab.

Education

2023 – 2025 🔖 **M.Sc. Artificial Intelligence** in Chonnam National University.
Thesis title: *Enhancing Polyp Detection and Segmentation in Medical Imaging via Deep Learning.*
GPA: 4.36/4.5.

2018 – 2022 🔖 **B.Sc. Control Engineering and Automation** in Hanoi University of Science and Technology.
Thesis title: *Advanced Deep Learning Network-based Medical Image Segmentation.*
GPA: 3.64/4.

Interest Research

2024 – Now 🔖 **Vision-Language for Medical Domain**
X-Ray Image and Report

🔖 **Multimodal Sentiment Analysis**
Sentiment Classification.

2023 – Now 🔖 **Mental Health Analysis**
Depression, Anxiety, Suicide Detection.

2021 – Now 🔖 **Medical Image Analysis**
Segmentation, Detection, Classification.

Research Publications

Conference Proceedings

1 V. T. H. Son, **N. Q. Vinh**, and S.-H. Kim, "Individual audio-driven talking head generation based on sequence of landmark," in *Annual Conference of KIPS 2024*, 2024.

2 **N. Q. Vinh**, V. T. H. Son, and S.-H. Kim, "Polyp-ses: Automatic polyp segmentation with self-enriched semantic model," in *17th Asian Conference on Computer Vision*, 2024. 🔗 URL: https://arxiv.org/pdf/2410.01210.

3 **N. Q. Vinh**, V. T. H. Son, V. H. C. Truong, D. D. Nguyen, M. N. H. Nguyen, and S.-H. Kim, "Rethinking top probability from multi-view for distracted driver behaviour localization," in *The IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshop*, 2024.

4 H.-C. Truong-Vinh, D.-K. Ta, D.-D. Nguyen, L.-T. Nguyen, and **N. Q. Vinh**, "Beyond keywords: Chatgpt's semantic understanding for enhanced media search," in *MediaEval 2023 Workshop*, CEUR-WS, 2023. 🔗 URL: https://ceur-ws.org/Vol-3658/paper6.pdf.

5 **N. Q. Vinh**, V. T. Huynh, and S.-H. Kim, "Adaptation of distinct semantics for uncertain areas in polyp segmentation," in *34th British Machine Vision Conference*, BMVA, 2023. 🔗 URL: https://papers.bmvc2023.org/0806.pdf.

6 **N. Q. Vinh**, T.-T. Tran, and V.-T. Pham, "Gca-net: Geometrical constraints-based advanced network for polyp segmentation," in *2022 9th NAFOSTED Conference on Information and Computer Science (NICS)*, 2022, pp. 241–246. 🔗 DOI: 10.1109/NICS56915.2022.10013367.

Under Review

1 **N. Q. Vinh**, M. D. Nguyen, T. H. S. Vo, H.-J. Yang, and S.-H. Kim, *Anatomical attention alignment representation for radiology report generation*, 2025.

2 **N. Q. Vinh**, T. D. Nguyen, D. D. Nguyen, *et al.*, *Time-aware mental state space for depression detection on social media*, 2025.

3 V. T. H. Son, **N. Q. Vinh**, and S.-H. Kim, *Kan-based fusion of dual-domain for audio-driven facial landmarks generation*, 2024.

4 V. T. H. Son, **N. Q. Vinh**, J.-e. Shin, S. Kim, H.-J. Yang, and S.-H. Kim, *Atl-diff: Audio-driven talking head generation using early landmark guide noise diffusion*, 2024.

Project

2024   🔖 **Depression Detection**
In this project, I developed a deep learning model for detecting depression from multi-modal data on social media. The model leverages an attention mechanism, a selective state space model, and transformers to effectively analyze both visual and textual information. The method achieves competitive results: a F1 score of 0.986 on Twitter and a F1 score of 0.950 Reddit datasets outperforming recent SOTA approaches on depression detection.

🔖 **Action Localization**
This project aims to recognize driver behaviors during real-world driving using video data from three cameras (dashboard, rear, right side). Drivers perform 16 distracting activities like drinking, phone calls, and texting, ... Our method uses the VideoMAE model for action recognition, combining multiview predictions to enhance accuracy. Conditional post-processing refines time boundaries and actions by leveraging output probabilities. The project achieved 6th place in the AI City Challenge at the CVPR Workshop 2024.

Project (continued)

Chatbot
This project involved developing a specialized math chatbot using a large language model. The system offers various advanced features, including conversational interaction similar to ChatGPT, answering complex math questions, interacting with PDF files, performing OCR on math equations, and visualizing mathematical expressions.

2023   **Tracked-vehicle Retrieval**
This project focuses on retrieving tracked vehicles based on natural language descriptions. We employed X-CLIP, a text-video contrastive learning module, to combine video and text data. A CLIP-based domain adaptation is integrated to establish vehicle color and type modules, which facilitate enhanced vehicle attribute matching. This approach enables more accurate retrieval of vehicles based on the given descriptions.

2022   **Event Retrieval**
In this project, I developed a system for retrieving events from visual data, focusing on TV news footage that encompasses a wide range of global and national events. The system utilizes language–image models like CLIP and BLIP to accurately retrieve events based on visual data. A user-friendly interface allows users to input text queries to efficiently explore visual information in real-time news contexts.

Face-Anti-Spoofing
This project aims to determine the authenticity of captured multi-modal face images, classifying them as real or fake. Our approach uses a multi-stream model to extract distinct features from each modality, which are then fused for more robust analysis. Our approach has demonstrated strong performance in distinguishing between genuine and spoofed images.

2021   **Smart Home**
In this project, I developed a smart home automation system using Google Assistant and the ESP8266 microcontroller. Users can control various home functions such as turning lights and fans on/off, closing doors, and monitoring temperature and humidity. The system supports seamless interaction, allowing control through either the web UI or voice commands via Google Assistant, providing an efficient and user-friendly smart home solution.

Skills

Languages	Fluent, with proficiency in technical writing and academic communication
Coding	Python, C, C++, LaTeX
Framework	Pytorch, Tensorflow, Transformer, LangChain, Hugging Face
Library	Numpy, Seaborn, Sklearn, SciPy, OpenCV
Web Dev	Html, Css, JavaScript
Misc.	Academic research, teaching, training, consultation, LaTeX typesetting and publishing
Hobby	Korean language, sport.

Miscellaneous Experience

Talks

2024　　■　**FPT Software Quy Nhon**, AI in Medical Image Analysis

Awards and Achievements

■　**Best Paper Award**, International Conference on Big Data Applications and Services

■　**Rank 6th,** Naturalistic Driving Action Recognition, AI City Challenge, CVPRW 2024

2023　　■　**Graduate School RA (Research Assistant) Scholarship**, Chonnam National University

■　**Chonnam National University Industry-Academic Cooperation Group**, Chonnam National University

■　**Global Scholarship**, Chonnam National University

2018　　■　**Pham Van Tra Schorlaship**, Bacninh Province

■　**Second Prize for Math Competition**, Bacninh Province

Certification

2024　　■　**Best Paper Award**. Awarded by Korea Big Data Service Society

References

Prof Soo-Hyung Kim
Professor
Chonnam National University,
77 Yongbong-ro, Buk-gu, Gwangju, South Korea
shkim@jnu.ac.kr

Assoc. Prof. Tran Thi Thao
Associate Professor
Hanoi University of Science and Tecnology,
No. 1, Dai Co Viet Road, Hanoi, Vietnam
thao.tranthi@hust.edu.vn

Assoc. Prof. Pham Van Truong
Associate Professor
Hanoi University of Science and Tecnology,
No. 1, Dai Co Viet Road, Hanoi, Vietnam
truong.phamvan@hust.edu.vn