Tri Cao

Email: tricao2001vn@gmail.com | Vietnamese/ Research Assistant at NUS | Homepage /Google Scholar

Education

School of Computing, National University of Singapore (NUS) 08/2025 (Expected)
Singapore City, Singapore
Intended Thesis Advisors: *Prof. Bryan HOOI* , *Prof. Shuicheng YAN*

University of Science, Vietnam National University (VNUHCM) 08/2019 - 08/2023
Ho Chi Minh City, Vietnam
Bachelor of Computer Science - Honors Program - GPA: 3.84/4.0 - Fully in English
Graduation Rating: Excellent (First class honours)
Thesis: Anomaly Detection under Distribution Shift. Mark: 10/10

Research Interests

Trustworthy AI, AI Agent, Anomaly Detection, Medical Image Analysis.

Publications

1. Tri Cao*, Chengyu Huang*, Yuexin Li*, Huilin Wang, Amy He, Nay Oo, Bryan Hooi. **PhishAgent: A Robust Multimodal Agent for Phishing Webpage Detection.** The Conference on Artificial Intelligence (AAAI) 2025 - Oral Presentation. *Link.*

2. Tri Cao, Jiawen Zhu, Guansong Pang. **Anomaly detection under Distribution Shift.** The International Conference on Computer Vision (ICCV) 2023. *Link.*

3. Ailin Deng, Tri Cao, Z. Chen, Bryan Hooi. **Words or Vision: Do Vision-Language Models Have Blind Faith in Text?** The Conference on Computer Vision and Pattern Recognition (CVPR) 2025. *Link.*

4. Yuexin Li, Hiok Kuek Tan, Qiaoran Meng, Mei Lin Lock, Tri Cao, Shumin Deng, Nay Oo, Hoon Wei Lim, Bryan Hooi. **PhishIntel: Toward Practical Deployment and Monitoring of Reference-based Phishing Detection.** The World Wide Web Conference (WWW) 2025 - Demo Paper. *Link.*

5. Yuexin Li, Chengyu Huang, Shumin Deng, Mei Lin Lock, Tri Cao, Nay Oo, Hoon Wei Lim, Bryan Hooi. **KnowPhish: Large Language Models Meet Multimodal Knowledge Graphs for Enhancing Reference-Based Phishing Detection.** USENIX Security 2024. *Link.*

6. Duy Minh Ho Nguyen, [et al., including Tri Cao and Binh T. Nguyen]. **LVM-Med: Learning Large-Scale Self-Supervised Vision Models for Medical Imaging via Second-order Graph Matching.** The Conference on Neural Information Processing Systems (NeurIPS) 2023. *Link.*

7. Duy Minh Ho Nguyen, [et al., including Tri Cao and Binh T. Nguyen]. **Joint Self-Supervised Image-Volume Representation Learning with Intra-Inter Contrastive Clustering.** The Conference on Artificial Intelligence (AAAI) 2023. *Link.*

Under Review

1. Tri Cao, Huy Trinh, Ailin Deng, Nam Nguyen, Khoa Duong, Man Cheung, Bryan Hooi. **Are Anomaly Scores Telling the Whole Story? A Benchmark for Multilevel Anomaly Detection**. *Link.*

2. Tri Cao, Bennett Lim, Yuexin Li, Shumin Deng, Yue Liu, Yuan Shui, Nay Oo, Shuicheng Yan, Bryan Hooi. **VPI-Bench: Visual Prompt Injection Attacks for Computer-Use Agents.**

Research Experience

National University of Singapore, Singapore - Research Assistant 01/2024 - Present
Supervisor: *Prof. Bryan HOOI*

Topic: Safety AI Agent.

First author of "VPI-Bench: Visual Prompt Injection Attacks for Computer-Use Agents." (Under Review).

- Identified attack strategies and created a new visual prompt injection dataset for computer-use agents.
- Evaluated state-of-the-art computer-use agents on the created dataset.
- Conducted vulnerability analysis of computer-use agents against prompt injection attacks.
- Primarily responsible for the manuscript.

Topic: Phishing Webpage Detection.

Co-first author of *"PhishAgent: A Robust Multimodal Agent for Phishing Webpage Detection"* (AAAI 2025).

- Constructed an agent that dynamically interacts with and manipulates tools, utilizing MLLMs for enhanced detection capabilities and robustness against adversarial attacks.
- Contributed to conducting a literature review about phishing detection models and selecting baseline models and datasets.
- Primarily responsible for the manuscript.

Deployment: *Worked with NUS IT to integrate PhishAgent into NUS email systems.*

*Equal contribution.

Topic: Multilevel Anomaly Detection

First author of *”Are Anomaly Scores Telling the Whole Story? A Benchmark for Multilevel Anomaly Detection”* (Under Review).

- Proposed a novel settings - Multilevel Anomaly Detection.
- Conducting comprehensive experiments using various models from different approaches to make the benchmark for the Multilevel Anomaly Detection setting.
- Conducted various findings and analyses based on the benchmark results.
- Primarily responsible for the manuscript.

Topic: Visual-Text Conflict in Large Vision-Language Model.

Second author of *”It Takes Two: Revealing Model Behavior under Vision-Language Data”* (CVPR 2025).

- Conducted experiments using various MLLM models on the MathVista and Brand Recognition datasets to investigate the behavior of the models in the context of text-visual conflicts.
- Participated in discussion about the findings from the experimental results.
- Involved in writing the manuscripts.

Singapore Management University, Singapore - Research Assistant 09/2022 - 03/2023
Supervisor: *Prof. Guansong PANG* - Topic: Anomaly Detection on image data.

First-author of *”Anomaly Detection under Distribution Shift”* (ICCV 2023).

- Proposed an Anomaly Detection method to address the distribution-shift challenge.
- Introduced a benchmark for anomaly detection under condition of distribution shift.
- Led the execution of experiments, which encompassed tasks such as baseline selection, dataset choice, method implementation, ablation study design, and comprehensive result analysis.
- Primarily responsible for the manuscript and rebuttal period.

German Research Center for AI, Germany - Remote Research Intern 03/2022 - 08/2023
Supervisor: *Prof. Binh NGUYEN*, *Duy NGUYEN*
Topic: Unsupervised Learning for 2D/3D Medical Image Processing.

Co-author of *”LVM-Med: Learning Large-Scale Self-Supervised Vision Models for Medical Imaging via Second-order Graph Matching”* (NeuRIPS 2023) and
Co-author of *”Joint Self-Supervised Image-Volume Representation Learning with Intra-Inter Contrastive Clustering”* (AAAI 2023).

- Participated in making literature review about self-supervised learning for medical imaging and the selection of baselines/datasets.
- Implemented and evaluated baseline models across multiple datasets.
- Executed a part of the experiments for downstream tasks.

Industrial Experience

Katalon,Inc - Vietnam - Applied Research Associate 09/2021 - 07/2022
Superviser: *Prof. Vu NGUYEN* - Topic: AI for Software Testing.

First-author of *”Ensemble approach for UI test case prioritization and selection”* (SEKE 2022).

- Ensemble Approach for Enhancing Robustness in UI Test Case Prioritization and Selection.
- Led the execution of experiments, which encompassed tasks such as baseline selection, dataset choice, method implementation and comprehensive result analysis.
- Primarily responsible for the manuscript.

Deployment: *Collaborated with the engineering team to integrate the method into Katalon’s testing software.*

Awards

Outstanding Student Research Award from University of Science (3 consecutive years). 2020-2023
Encouragement Scholarships for outstanding students from University of Science (5 times). 2020-2023
Merit Certificate from Vietnam National University Director for outstanding graduates. 2023
Fully-funded Bachelor Scholarship from University of Science, Vietnam. 2021-2022
Listed on the **”Gold Board”** for achieving a top GPA at University of Science. 2021-2022
Odon Vallet Scholarship for outstanding undergraduate students, France. 2022

Competitions

SHREC AI Challenge: **2nd place** in Sketch-Based 3D Shape Retrieval in the Wild track. 2022
MediaEval AI Challenge: **1st place** in NewsImage task and **2nd place** in FakeNews task. 2021

Professional Activities

Reviewer at ICML, ICLR, CVPR, ICCV 2024, 2025
Mentor at Math and Science Summer Program (MaSSP), Vietnam 07/2024
Speaker at DS@UIT Winter Workshop, Ho Chi Minh City, Vietnam 10/2023
Selected students participating in Summer School at HUST, Ha Noi. 07/2023
Selected outstanding student participating in the Meeting and Dialogue Program 04/2023
between city leaders and outstanding students from universities in Ho Chi Minh city.