MINGYUAN LI

myuli@seas.upenn.edu

mingyuanli.site

EDUCATION AND AWORDS

University of Pennsylvania, *Philadelphia, PA, USA* Aug./2025-Now

Master of Science in Engineering (MSE), Systems Engineering

Xi'an Jiaotong-Liverpool University, *Suzhou, China* Sept./2021-Jul./2025

Information and Computing Science *(Bachelor of Science with Honours (Class I))* **WES GPA** 3.82/4, **Major GPA** 3.87/4

Won the Excellent Poster Prize and School Winner in SURF 2024 ***Top 5%*** Oct./2024
Won the Global Impact Grants 2023-24 (**£1000**) from Advance HE ***Top 15*** May/2024
Won the 3rd Prize (**3000RMB**) in the IEEE CyberC 2023 Data Analysis Competition ***Top 3*** Nov./2023
Won the 2nd Award (**1000RMB**) at the 2023 XJTLU Student Research-Oriented Learning Summit Nov./2023
Won the Honorable Mention for MCM/ICM ***Top 21%*** May/2023

PUBLICATIONS

Mingyuan Li; Duan Wang; Erick Purwanto; Hai-Ning Liang, **"Understanding Data Containers with Multimodal LLMs: An Empirical Study Across Encoding Formats and Visual Data"**, Under review at ***IEEE Transactions on Software Engineering (TSE)*** (Submitted Nov. 2025).

Mingyuan Li; Duan Wang; Erick Purwanto; Thomas Selig; Qing Zhang; Hai-Ning Liang, **"VisualCodeMOOC: A Course Platform for Algorithms and Data Structures Integrating a Conversational Agent for Enhanced Learning Through Dynamic Visualizations"**, ***SoftwareX***, vol. 30, p. 102072, May 2025, doi: 10.1016/j.softx.2025.102072

Erick Purwanto; Na Li; Qing Zhang; Thomas Selig; Yihong Wang; Teng Ma; Filbert Juwono; Pengfei Fan, (student member) **Mingyuan Li** and Duan Wang, **"Chatbot-Powered Learning for Sustainable Education in Programming"**, advance-he.ac.uk/knowledge-hub/global-impact-grants-2024-education-sustainable-development-and-building-future

Mingyuan Li; Yichuan Wang; Junfeng Huang; Erick Purwanto; Ka Lok Man, **"Patch-Based Multi-Level Attention Mechanism for Few-Shot Multi-Label Medical Image Classification"**, ***The 15th International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery***, pp. 84-91, 2023, doi: 10.1109/CyberC58899.2023.00024

Doyeon Kim; Xujia Ning; Kaicheng Liang; Yi Ni; Duan Wang; **Mingyuan Li**; Yichuan Wang; Erick Purwanto; Ka Lok Man, **"Cervical Spine Fracture Detection Through Two-Stage Approach of Mask Segmentation and Windowing Based on Convolutional Neural Network"**, ***International Conference on Platform Technology and Service (PlatCon)***, pp. 1-6, 2023, doi: 10.1109/PlatCon60102.2023.10255157

RESEARCH EXPERIENCES

Exploring Equity–Vulnerability Balances in Structured Visual AI for Learning Feb./2025-Now

Supervisor**: Dr. Yihong Wang; Dr. Erick Purwanto; **Lead Research Assistant

- Performed data cleaning and feature engineering on the OvSV dataset derived from 8 controlled experiments with 330 students to enable statistical modeling.
- Conducted EFA and removed 9 features; validated a 5-factor model using CFA (TLI/CFI > 0.90; RMSEA/SRMR < 0.08), ensuring questionnaire construct quality.
- Based on the CFA model, assessed Cronbach's α and Pearson correlations, performed MANOVA and t-tests, and used ANOVA to analyze between-group differences across features, revealing 9 key significant effects.
- (In progress) Authoring a manuscript (literature review, methodology, findings, discussion) for *British Journal of Educational Technology*.

Exploring LLM's Impact on Software Engineering via Encoding Formats and Visual Data Dec./2024-Nov./2025

Supervisor:** Dr. Erick Purwanto; Dr. Hai-Ning Liang; **Research Assistant (Team of Two)

- Iterated VisualCodeChat to ver. 2.0, incorporating multimodal capabilities.
- Constructed a systematic dataset comprising 300 encoding elements and collected 1500 dialogue instances via it.
- Designed and conducted static analysis, open card-sorting discussions, and comprehensive quantitative analyses to evaluate LLMs' capabilities in generating encoding formats for algorithmic formulation (RQ1).
- Developed 3 fixing prompt strategies to assess LLMs' capabilities in repairing and iteratively repairing errors identified in RQ1.
- Authored a manuscript (Introduction, background, study setup, RQ1, RQ2) currently under review at *IEEE TSE*.

GenAI-Powered ChatBot in Programming and Algorithm Education Sept./2023-Dec./2024

Supervisor**: Dr. Erick Purwanto; Dr. Yihong Wang; Dr. Thomas Selig; **Research Assistant (Team of Two)

Phase 1: Created a programming tutor bot based on open-source projects for programming and algorithm

- Utilized Prompt Engineering to call OpenAI APIs, creating an educational chatbot with eight different topics.

- Implemented Chain of Thought and Moderation designed language interaction pipeline, customized for beginners.
- Test our bot with an Indonesian student across four topics with two languages, achieving a 91% effectiveness rate, demonstrating its pedagogical value and multilingual capabilities.

Phase 2: Developed algorithm visualization modules based on bot responses to aid teaching.

- Created 7 algorithm visualizations using d3.node, including for loops and array sorting lessons.
- Refactored the backend logic of a React-based TypeScript project to synchronize bot responses with dynamic visualizations, leading to the development of VisualCodeChat, a teaching-focused chatbot platform.
- Developed a pilot MOOC with VisualCodeChat and further integrated our chatbot into an existing mature platform to create VisualCodeMOOC.
- Conducted 2 controlled experiments consecutively with 16 preA high school students and 88 non-programming background students from the CPT206 course. Reliable questionnaire results (Cronbach's alpha: 0.891) and a high average score of 4.18, supported by qualitative coding, demonstrate our bot met all evaluation criteria.

Phase 3: Researched the effectiveness of our design compared to standard GPT in teaching graph algorithms.

- Iterated 3 graph algorithm visualizations using d3. force, including DFS-based cycle and connectivity checks, and BFS-based connected components check; added data recording and user-friendly guidance modules.
- Conducted 8 controlled experiments (330 CPT204 students) to evaluate effectiveness, producing the OvSV dataset with two groups: our bot vs. a original GPT-based AI tool.
- Our solution outperforms ChatGPT across all dimensions, with T-test p-values below 0.05, confirming statistically significant improvements in usability, effectiveness, and engagement.

Few-Shot Multi-Label Medical Image Classification Research Jul.-Sept./2023

Supervisor**: Dr. Erick Purwanto;* ***Research Fellow

- Trained a few-shot classification model using the VPT with a backbone of Swin-transformer.
- Evaluated robustness using 3 datasets: ChestDR (chest X-ray), Endo (genuine colonoscopy), and Colon (colon cells).
- Developed multi-level attention patch-based preprocessing technique for the model, enhancing the model's ability to detect minute details, e.g., overlapping information.
- Improved mAP (1.2%-1.7%) and AUC (4.1%-5.2%) on two datasets (ChestDR, and Endo), compared to the baseline.

Replication and Improvement of the Cervical Spine Fracture Detection Research Jun.-Jul./2023

Supervisor**: Dr. Erick Purwanto;* ***Research Fellow

- Developed a two-stage approach for cervical spine CT scan analysis, achieving a combined accuracy of 94.9%, and a BCE logits coefficient of 0.20 ± 0.01.
- Utilized UNet-EfficientNet in Stage 1 for precise CT image segmentation, attaining an accuracy of 99.91%.
- Applied CrackNet-LSTM in Stage 2 for accurate cervical spine fracture detection, achieving an accuracy of 94.9%.

COURSEWORK, PRESENTATION AND SEMINAR

Presented SURF research results: VisualCodeChat: Dynamic Programming Tutor with Visual and Personalized Feedback in 2024 SURF Poster Fair Oct./2024

Introduction to Computer Networking: Developed a client application for user authorization and file transfer using Python Sockets; simulated SDN network traffic control with Mininet. Oct.-Dec./2023

Software Engineering Group Project: Led an 8-member team to develop a sports center booking system based on the MVC, completing full-stack development of two admin modules. Mar.-May/2024

Human-Centric Computing: Collaborated on requirement analysis, design, prototyping, and evaluation in a team, leading to a presentation, report, poster, and an Honorable Mention (1 of 8). Mar.-May/2024

Advanced OO Programming: Developed a BFS-based multi-agent pathfinding game with a graphical interface, demonstrating algorithm superiority through interactive validation. Mar.-May/2024

INTERNSHIP

Guangzhou Hehui Technology Co. *Java Software Engineer Intern* Mar.-Jun./2023

***Supervisor**: Xiaolan Zhou*

- Contributed to the development of a pharmaceutical ordering system.
- Optimized the code structure with a low coupling design principle to ensure maintainability and scalability.
- Effectively managed MySQL databases, and designed database structures that adhere to the Third Normal Form.

PROFESSIONAL SKILLS

Programming Languages: Java (3yrs+), MySQL (3yrs+), Python (2yrs+), C++ (1yr+), TypeScript (1yr+), C# (1yr+)

Framework: React (1yr+), Next.js(1yr+), Spring Boot (1yr+)

Library: MMCV, OpenCV, PyCOCO, ultralytics, PyTorch, scikit-learn, D3.js