

# Xu (Oscar) Li

Tel: +1 617 8746480 | Email: [li.xu2@northeastern.edu](mailto:li.xu2@northeastern.edu) | <https://xuli0929.github.io>

## EDUCATION

**Northeastern University**, Boston, MA

May, 2026 (Expected)

Bachelor of Science in Computer Science

GPA: 4.00

Major Courses: Machine Learning and Data Mining, Artificial Intelligence, Algorithms and Data, Probability and Statistic, Mathematics of Data Models, Object-Oriented Design

## Publication

1. Li, X., & Lyu, F. (2025). MM-Prompt: Cross-Modal Prompt Tuning for Continual Visual Question Answering. *arXiv preprint arXiv:2505.19455*. (Submitted, Under Review)
2. Li, X. (2025). Enhancing earthquake detection systems through AI-based frameworks for multimodal learning and sensor integration models. *DYNA*. (Accepted, April 2025).
3. Li, X. (2025). Enhancing video-based emotion recognition with multi-head attention and modality dropout. In *Proceedings of CONF-MLA. EAI*.
4. Li, X. (2024). Exploring the effect of depth and width of CNN models on binary classification of dogs and cats. *Applied and Computational Engineering*.

## Research Experience

**UC Berkeley**

Berkeley, United States

*Visiting Student Researcher — AI4EPS supervised by Prof. Weiqiang Zhu*

May 2025 - Present

Project: First AI Foundation Model in Earth and Planetary Science (Ongoing)

- Developing specialized encoders for seismic waveforms as existing audio/vision models fail to capture critical geophysical characteristics and poor geological universality.
- Adapting self-supervised learning paradigms (AR, Wav2Vec2, MAE) with geophysics-informed innovations: station distance bias correction, wave-physics-aware feature extraction.
- Future Direction: Integrating seismic encoders with language models, enabling natural language queries for earthquake analysis and automated interpretation of geophysical phenomena.

**Collaborative Researcher**

Sep 2024 - May 2025

*With Dr. Fan Lyu, Postdoc, Institute of Automation, Chinese Academy of Sciences*

Project: MM-Prompt: Cross-Modal Prompt Tuning for Continual Visual Question Answering

- Existing prompt-based CVQA methods process the prompts for each modality independently, which provide model with more already-preferred information and amplifies modality imbalance, leading to degraded multi-modal performance as tasks accumulate.
- Developed MM-Prompt framework with two novel components: cross-modal query enriching prompts with opposite-modality features before selection, and cross-modal recovery using hierarchical masking-reconstruction to embed fused knowledge into prompt space.
- Achieved state-of-the-art performance on VQA v2 and NExT QA, demonstrating superior accuracy retention and balanced modality utilization across sequential tasks compared to prior prompt-based methods that suffer from progressive modality bias.

**Northeastern University**

Boston, United States

*Undergraduate Researcher — ChatsLab supervised by Prof. Weiyang Shi*

Dec 2024 - Present

Project 1: Safety in GUI Agent under Multi-Round Conversation (Ongoing)

- Identified critical gap in GUI agent safety: while single-round jailbreaks are well-studied, multi-round attacks requiring complex context understanding and selective task execution remain unexplored.
- Developed automated multi-round evaluation framework extending single-turn GUI agents, implementing round transition logic and LLM-based validators to assess agents' ability to maintain safety boundaries across conversation history.
- Created comprehensive safety benchmark with three key metrics: defend rate measuring successful refusals of harmful tasks, jailbreak rate tracking harmful task completions, and safe successful rate quantifying legitimate task completions, enabling evaluation of agents' ability to distinguish and selectively execute safe vs. harmful instructions.

Project 2: LLM Persuasion Arena

- Built comprehensive benchmark platform evaluating LLM persuasion capabilities across multiple models through controlled scenarios (donation, debate, discussion) with configurable techniques (evidence-based, shared values).
- Designed persona-driven conversation framework with fine-grained agent control (demographics, personality, beliefs), enabling systematic testing of persuasion strategies against diverse synthetic populations.
- Developed dual-benchmark system for evaluating LLMs as personalized judges: both models predict conversation outcomes, but while one receives full user personas, the blind model must additionally infer these characteristics. Testing whether LLMs can simultaneously understand behavioral patterns and make accurate personalized predictions.

**University of Electronic Science and Technology of China (UESTC)** Chengdu, China  
*Visiting Student Researcher — International Lab of Intelligent Information Processing* Jun - Dec, 2024

**Project:** Multi-Modal Model for Earthquake Prediction

- Developed multi-modal AI framework integrating seismic data (waveforms, drill data, electromagnetic signals) for enhanced earthquake detection and early warning.
- Designed adaptive modality-weighting mechanism using attention-based reliability scoring to dynamically prioritize data sources based on signal quality and geological context, maintaining robustness under sensor failures.
- Implemented cross-modal fusion strategies combining spatial-temporal features with physics-informed constraints, leveraging inter-modal correlations to capture precursor patterns.

## PROJECTS EXPERIENCE

---

**Vision and Insight**

Feb - Apr, 2025

- Applied interpretability tools (Grad-CAM, filter activation) to diagnose CNN weaknesses in art style classification, leveraging insights to integrate CBAM and SPP/CIM modules, improving accuracy from 68.64% to 78.68%.

**Algorithms for Big Data Models and Applications**

Remote

*Team Member — Mentor: Prof. David.P. Woodruff, Carnegie Mellon University* Aug - Oct, 2024

- Conducted comprehensive literature review on landmark recognition/retrieval, identifying critical research gaps in robustness to occlusion and viewpoint variation, providing strategic insights for future model development in CNN and hybrid architectures.

**Multi-Dimensional Applications of Machine Learning and Deep Learning**

Remote

*Team Leader — Mentor: Prof. Bjorn Schuller, Imperial College London* Jul, 2024 - Sept, 2024

- Designed CFNSR-MAFNet, a multi-modal emotion recognition model with cross-modal fusion, modality dropout, and residual connections, achieving 78.3% accuracy on RAVDESS dataset by effectively combining audio and video features for robust performance.

**AI Deep Learning and Causality**

Remote

*Team Member — Mentor: Prof. Nicholas Lane, University of Cambridge* Jul - Aug, 2023

- Optimized CNN architectures for large-scale image classification, achieving 95.4% accuracy through systematic analysis of network depth and width trade-offs.

## SKILLS

---

- Technical Skills: Python, Java, R, Kotlin, SQL.
- Languages: Chinese (Native), English (Advanced, barrier-free business communication), Japanese (Limited working proficiency)
- Test Score: GRE: 334