Yiyang Huang

◇ huang.yiyan@northeastern.edu ◇ Personal Page ◇ Google Scholar ◇ LinkedIn

Education

Sep 2024 – Present	**Northeastern University, USA** Ph.D. in Computer Engineering, advised by Prof. Yun Raymond Fu
Sep 2021 – Jun 2024	**Xidian University, China** M.S. in Computer Science, advised by Prof. Xuefeng Liang
Sep 2017 – Jun 2021	**Xidian University, China** B.Eng. in Intelligence Science and Technology

Research Interests

Multimodal LLMs; Efficiency; Reliability; Hallucination Detection & Mitigation; Video Understanding; Layout Understanding

Experience

Northeastern University **Sep 2024 – Present**
SMILE Lab, Research Assistant | Supervisor: Prof. Yun Raymond Fu **Boston, USA**

- Published **1 ICLR paper** on mitigating hallucinations in MLLMs via encoder-level inference intervention.
- Published **1 EMNLP (Main) paper** on scaling image-pretrained MLLMs to long-video understanding.
- Identified three encoder-side causes of hallucinations (statistical bias, inherent bias, vulnerability) and developed SHIELD, a plug-and-play token-editing module (re-weighting/subtraction + adversarial contrastive decoding) to suppress hallucinations in captioning and VQA.
- Developed D-CoDe, a plug-and-play pipeline that decomposes questions into targeted sub-queries and performs dynamic visual token/frame compression to enable long-horizon video QA under tight context budgets.
- Authored a survey on Video-LLM hallucinations and maintained a curated GitHub repository with a structured taxonomy and benchmark/metric/method tables. [GitHub]

Adobe Research **May 2025 – Nov 2025**
Vision-Language Lab, Research Intern | Mentors: Zhaowen Wang, Simon Jenni, Jing Shi **San Jose, USA**

- Submitted **1 CVPR paper** and filed **1 patent** on compositional layout understanding in VLMs.
- Diagnosed two failure modes in layered designs (semantic drift and structural ambiguity) and built MASON, a plug-and-play framework that adds metadata-aware multimodal alignment and structural cue injection for compositional reasoning.
- Built the CoDeLayout dataset (~20K instances) with an automated pipeline for compositional pair mining and QA generation.

Xidian University **Sep 2021 – Jun 2024**
Research Assistant | Supervisor: Prof. Xuefeng Liang **Xi'an, China**

- Published **1 ACM MM paper** on contrastive and attribute learning for visual speech recognition.
- Published **1 ICASSP paper** on visual speech recognition in low-resource language settings.
- Proposed CALLip, a multimodal framework that leverages attribute learning to normalize cross-speaker lip shape variations and audio-visual contrastive learning to mitigate viseme confusion.
- Developed dynamic LoRA to adapt visual encoders for learning meta lip shapes shared across languages, and applied multilingual instruction tuning with large language models to improve cross-lingual text mapping in low-resource settings.

Kyoto University **Sep 2023 – Mar 2024**
Research Student | Mentor: Prof. Takatsune Kumada **Kyoto, Japan**

- Developed a social-scene MLLM framework that integrates body language cues via multiple visual branches.
- Applied Chain-of-Thought reasoning guided by cognitive theory to infer relationships among nonverbal cues, improving few-shot in-context learning and reducing modal bias and hallucinations in social interaction understanding.

Publications

Under Review / Preprint

- **Yiyang Huang**, Yitian Zhang, Yizhou Wang, Mingyuan Zhang, Liang Shi, Huimin Zeng, Yun Fu. "Distorted or Fabricated? A Survey on Hallucination in Video LLMs." ARR under-review.
- **Yiyang Huang**, Zhaowen Wang, Simon Jenni, Jing Shi, Yun Fu. "MASON: Compositional Design Layout Understanding in VLMs through Multimodal Alignment and Structural Perception." CVPR under-review.

Conferences

- **Yiyang Huang**, Liang Shi, Yitian Zhang, Yi Xu, Yun Fu. "SHIELD: Suppressing Hallucinations in LVLM Encoders via Bias and

Vulnerability Defense." ICLR 2026. [Paper] [Code]
- **Yiyang Huang**, Yizhou Wang, Yun Fu. "D-CoDe: Scaling Image-Pretrained VLMs to Video via Dynamic Compression and Question Decomposition." EMNLP 2025 (Main). [Paper] [Code]
- Shuai Zou, Xuefeng Liang, **Yiyang Huang**. "LipReading for Low-resource Languages by Language Dynamic LoRA." ICASSP 2025. [Paper]
- **Yiyang Huang**, Xuefeng Liang, Chaowei Fang. "CALLip: Lipreading using Contrastive and Attribute Learning." ACMMM 2021. [Paper]

Academic Service

Conference Reviewer: ARR, FG
Journal Reviewer: ACM Transactions on Knowledge Discovery from Data (TKDD)

Skills

Languages: Python, C++, MATLAB, LaTeX
Libraries: PyTorch, PyTorch Lightning, NumPy, Scikit-learn, Skimage

Awards

- Outstanding Student, Xidian University, 2022
- National Scholarship, China, 2021
- Undergraduate Computer Design Competition (1st Prize), China, 2021
- RoboMaster National Robotics Competition (2nd Prize), China, 2019
- ICRA AI Challenge (3rd Prize), 2019