Yiyang Huang

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

Education

Northeastern University, Boston, USA **Sep 2024 – Present**
Ph.D. in Computer Engineering, advised by Prof. Yun Raymond Fu

Xidian University, Xi'an, China **Sep 2021 – Jun 2024**
M.S. in Computer Science, advised by Prof. Xuefeng Liang

Xidian University, Xi'an, China **Sep 2017 – Jun 2021**
B.Eng. in Intelligence Science and Technology

Research Interests

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

Experience

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

- Published **1 ICLR 2026 paper** on mitigating hallucinations in MLLMs via encoder-level inference intervention.
- Identified encoder-side causes of hallucinations and proposed SHIELD, a training-free module (token reweighting/subtraction + adversarial contrastive decoding), reducing hallucinations by 27% on captioning across MLLM families.
- Published **1 EMNLP 2025 (Main) paper** on scaling image-pretrained MLLMs to video understanding.
- Developed D-CoDe, a plug-and-play framework that scales image-pretrained MLLMs to VideoQA via question decomposition and dynamic token/frame compression, improving first-person video understanding accuracy by 29% under LLaVA-NeXT.
- Authored a comprehensive survey on Video-LLM hallucinations and maintain a curated GitHub repository tracking related papers and benchmarks. [GitHub]

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

- Submitted **1 CVPR paper** and filed **1 patent** on compositional layout understanding in MLLMs.
- Identified semantic drift and structural ambiguity as key challenges in compositional layout reasoning, and developed MASON with metadata-visual grounding and 3D structural perception, improving compositional matching by 26% over Gemini-2.5-Pro and 15% over GPT-5/GPT-o3 under a Qwen2.5-VL backbone.
- Built CoDeLayout (~20K samples) through an automated pipeline for compositional layout mining and QA generation.

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

- Published **1 ICASSP 2025 paper** on visual speech recognition in low-resource language settings.
- Developed a dynamic LoRA framework to learn meta lip representations shared across languages, improving accuracy by 15% on low-resource languages (e.g., Portuguese and Italian) using pretrained 3DCNN backbones and the multilingual LLM BLOOMZ.
- Published **1 ACM MM 2021 paper** on contrastive and attribute learning for visual speech recognition.
- Proposed CALLip, integrating attribute learning with audio-visual contrastive learning to normalize cross-speaker lip shape variations, reducing word error rate by 50% on multilingual benchmarks.

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

- Developed a social-scene MLLM framework integrating body language cues via multiple visual branches, and applied CoT reasoning guided by cognitive theory to reduce modal bias and hallucinations in social interaction understanding.

Publications

Accepted

- **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." ACM MM 2021. [Paper]

Submitted / Under Review

- **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.
- Mingyuan Zhang, Yue Bai, Yifan Wang, **Yiyang Huang**, Yun Fu. "Rethinking Fine-Tuning: Unlocking Hidden Capabilities in Vision-Language Models." CVPR under-review. [Paper]
- Liang Shi, **Yiyang Huang**, Yun Fu. "Capturing Individual Differences of Facial Expression for Authentic Expression Generation." FG under-review.

Academic Service

Conference Reviewing: ACL Rolling Review (ARR); IEEE Face & Gesture (FG)
Journal Reviewing: ACM TKDD

Teaching

Teaching Assistant

DS 5110: Essentials of Data Science — Fall 2025
DS 5020: Fundamentals of Linear Algebra and Probability — Spring 2026

Skills

Languages: Python, LaTeX
Frameworks/Tools: PyTorch, HF Transformers, HF Datasets, ffmpeg, PyAV, decord, Git, Slurm
Training/Finetuning: Accelerate, PEFT, LoRA/QLoRA, bitsandbytes
Inference: vLLM, FlashAttention
Evaluation: lmms-eval, POPE, CHAIR, MME

Awards

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