Xiangyu Zeng

✉ XiangyuZeng2001@Outlook.com · ☎ +86 15264367990

⚭ https://lanxingxuan.github.io/ · Google Scholar Link

Education Background

Shandong University   Sep. 2020 – Jun. 2024

B.Eng.   School of Computer Science and Technology

Course GPA: 92.1/100   Class Rank: 3/52   Comprehensive GPA: 102.1/110   Comprehensive Rank: 1/52

- *Recipient of National Scholarship, Shandong University Dean's Scholarship, First-Class Academic Scholarship (×2), and Outstanding Graduate of Shandong University*

Nanjing University   Sep. 2024 – Present

Ph.D. Candidate   School of Computer Science and Technology

- *Recipient of National Scholarship, Nanjing University Presidential Special Scholarship, First-Class Academic Scholarship*

Competition Experience

- 23rd CCF CSP Certification: 370 points (Top 1%)   Sep. 19, 2021
- ACM ICPC Asia Regional Contest: Silver Medal   Nov. 14, 2021
- CCSP (East China): Bronze Medal   Dec. 13, 2021
- ACM ICPC Asia Regional Contest: Bronze Medal   Apr. 3, 2022
- CUMCM (Shandong): First Prize   Sep. 18, 2022

Research Experience

Selected Works

- **StreamForest: Efficient Online Video Understanding with Persistent Event Memory**
  - Accepted by **NeurIPS 2025**, first author (co-first rank1), Spotlight (Top 3%)
  - *Developed StreamForest, a novel architecture for real-time streaming video understanding, featuring a Persistent Event Memory Forest for efficient long-term memory and a Fine-grained Spatiotemporal Window for enhanced short-term perception.*
  - *Achieved state-of-the-art performance across multiple benchmarks while maintaining high accuracy under extreme compression.*

- **TimeSuite: Improving MLLMs for Long Video Understanding via Grounded Tuning**
  - Accepted by **ICLR 2025**, first author
  - *Without relying on any external expert, TimeSuite can achieve expert-level performance in grounding tasks while maintaining considerable generalization QA capability and strong zero-shot capabilities.*
  - *The introduction of grounding tasks enhances the comprehensive understanding of long videos. We validated the feasibility of enhancing MLLM's comprehensive capabilities by integrating expert tasks.*

- **Adaptive Edge-Aware Semantic Interaction Network for Salient Object Detection in Optical Remote Sensing Images**
  - Accepted by **TGRS 2023**, first author
  - *Developed AESINet, a novel edge-aware semantic interaction network for salient object detection in optical remote sensing images, incorporating LDAM for unsupervised edge enhancement, MFEM for handling varying object scales, and DSIM for robust detection in cluttered scenes.*
  - *Achieved superior performance over 14 state-of-the-art methods across three benchmark datasets.*

Other Works

- **Online Video Understanding: OVBench and VideoChat-Online**
  - Accepted by **CVPR 2025**

- **Task Preference Optimization: Improving Multimodal Large Language Models with Vision Task Alignment**
  - Accepted by **CVPR 2025**

- **Make Your Training Flexible: Towards Deployment-Efficient Video Models**
  - Accepted by **ICCV 2025**

- **VTTS: Visual Test-Time Scaling to Reinforce Multimodal Reasoning by Iterative Perception**
  - Accepted by **NeurIPS 2025**

- **InternVideo2.5: Empowering Video MLLMs with Long and Rich Context Modeling**
  - Technical Report

- **VideoChat-Flash: Hierarchical Compression for Long-Context Video Modeling**
  - Undergoing Review, Github **400+ Star**

- **VideoChat-R1: Enhancing Spatio-Temporal Perception via Reinforcement Fine-Tuning**
  - Undergoing Review

- **Learning Goal-Oriented Language-Guided Navigation with Self-Improving Demonstrations at Scale**
  - Undergoing Review

- **DAG: Unleash the Potential of Diffusion Model for Open-Vocabulary 3D Affordance Grounding**
  - Undergoing Review

- **HOID-R1: Reinforcement Learning for Open-World Human-Object Interaction Detection Reasoning with Multimodal Large Language Model**
  - Undergoing Review

- **FreeRet: MLLMs as Training-Free Retrievers**
  - Undergoing Review

- **UniFlow: A Unified Pixel Flow Tokenizer for Visual Understanding and Generation**
  - Undergoing Review

*Last updated: **October 24, 2025***