

# Linxiang Peng

School of Artificial Intelligence and Data Science, University of Science and Technology of China

plx@mail.ustc.edu.cn | GitHub: flying05 | Homepage

## EDUCATION

---

**University of Science and Technology of China (USTC)**

**Sept 2023 – Present**

- **B.Eng. in Data Science** *Expected Jul 2027*
- **GPA: 3.7/4.3** | **Rank: 10/55**
- **Relevant Coursework:**
  - **Core AI & Signal Processing:** Machine Learning
  - **Mathematics & Theory:** Mathematical Analysis (93), Linear Algebra (91), Stochastic Processes (96), Complex Analysis B (96), Partial Differential Equations B (96)

## PUBLICATIONS

---

### **FIS-OT: Feature-Induced Optimal Transport for Unsupervised Action Segmentation**

*Accepted to IEEE International Conference on Multimedia and Expo (ICME 2026, CCF-B)*

*Linxiang Peng\**, *Xinyao Qin\**, *Jinhan Li*, *Di Yang*, *Jiangtao Wang*

## RESEARCH EXPERIENCE

---

### **Unsupervised Action Segmentation**

**July 2025 – March 2026**

*Independent Research Project, School of AI & Data Science, USTC*

*Mentored by Prof. Jiangtao Wang*

My research focuses on unsupervised action segmentation, which aims to partition untrimmed long videos into consecutive action segments, addressing the high cost of manual annotation.

- For RGB-based long videos, We introduced a triplet-based contrastive learning module to alleviate the over-reliance on early-stage noisy pseudo-labels in previous SOTA methods. By integrating it with optimal transport, my approach achieved state-of-the-art performance across multiple benchmarks.

*Accepted to ICME 2026(First author)*

- For skeleton-based sequence inputs, We extended the joint-decoupling paradigm used in prior works with a masking strategy and explicit physical constraints on reconstructed skeletons, further obtaining comprehensive SOTA results on various datasets.

*Submitted to ACM MM 2026(Co-first author)*

### **Visual Token Pruning for Multimodal Large Language Models**

**2025 – 2026**

*Research Project, School of AI & Data Science, USTC*

*Mentored by Prof. Jiangtao Wang*

Focused on visual token pruning for MLLMs to alleviate the quadratic attention complexity of massive visual tokens, while preserving multimodal reasoning capabilities.

- Uncovered the ineffectiveness of vanilla attention scores for MLLM token importance estimation, attributed to unimodal visual self-attention and Transformer attention sink bias.
- Proposed a two-stage token pruning framework, involving coarse semantic filtering in vision encoders and entropy-gated refinement in LLMs to suppress attention sink noise.
- Validated the framework on mainstream MLLMs across multiple image and video VQA benchmarks, achieving state-of-the-art performance under various pruning ratios.
- Realized significant efficiency gains in prefill speed, end-to-end inference and KV cache compression with minimal visual token retention.

*Submitted to ECCV 2026 (Second author)*

## AWARDS

---

- **China Undergraduate Mathematical Contest, Provincial First Prize** Dec 2024
- **MCM/ICM, Honorable Mention (H Award)** 2025
- **USTC Progress Scholarship (from 3.45/4.3 to 3.58/4.3)** 2023–2024
- **USTC Outstanding Student Scholarship, Silver Award (Top 10%)** 2024–2025

## SKILLS

---

- **Programming:** C, Python, LaTeX, SQL
- **Deep Learning:** PyTorch

## LANGUAGES

---

- **Chinese:** Native
- **English:** Proficient in reading and writing scientific papers