

# Jiayu Li

✉ jiayu320.li@connect.polyu.hk    ☎ (86)18920305219    🔗 <https://jiayu320.github.io/>    🌐 Jiayu320

## Education

---

### Hong Kong Polytechnic University

Sept 2024 – July 2026

MS in Information Technology

- GPA: 3.71/4.3 (Graduating with **Distinction**)
- **Coursework:** Natural Language Processing; Machine Learning and Data Analysis; Computer Vision and Image Processing etc.
- **Thesis:** Research on efficient collaboration between edge and cloud LLMs

### Hefei University of Technology

Sept 2020 – June 2024

BS in Mathematics and Applied Mathematics

- GPA: 3.35/4.3
- **Coursework:** Analytic Geometry; Discrete Mathematics; Numerical Analysis etc.
- **Thesis:** Research on Audio-Driven Head Motion Video Generation

## Publications

---

\* Equal Contribution / Joint First Author

### CiteVQA: Benchmarking Evidence Attribution for Trustworthy Document Intelligence

2026

Dongsheng Ma\*, **Jiayu Li**\*, Zhengren Wang\*, Yijie Wang, Jiahao Kong, Weijun Zeng, Jutao Xiao, Jie Yang, Wentao Zhang, Bin Wang, Conghui He

Submitted to **NeurIPS 2026** (Datasets and Benchmarks Track).

### HybridFlow: Adaptive Task Scheduling for Fast and Token-Efficient LLM Inference in Edge-Cloud Collaboration

2026

Jiangwen Dong, **Jiayu Li**, Tianhang Zheng, Wanyu Lin.

Accepted by **ICML 2026** (International Conference on Machine Learning).

### AgenticOCR: Parsing Only What You Need for Efficient Retrieval-Augmented Generation

2026

Zhengren Wang\*, Dongsheng Ma\*, Huaping Zhong\*, **Jiayu Li**, Yijie Wang, Tianyao He, Wentao Zhang, Bin Wang, Conghui He

Submitted to **NeurIPS 2026**.

### AscendKernelGen: LLM-Driven Kernel Generation for NPUs

2026

Xinzi Cao\*, Jianyang Zhai\*, Pengfei Li\*, Zhiheng Hu\*, ..., **Jiayu Li**, et al.

Accepted by **ACL 2026 Findings**.

## Invention Patent

---

### Instruction-controlled 3D facial animation generation method based on single image

2024

Yan Xing, **Jiayu Li** (First Student Author), et al.

China Patent Application No. 202410912502.5.

## Research Internships

---

### Shanghai Artificial Intelligence Laboratory

Dec 2025 – Present

Intern — Advisor: Bin Wang

- **CiteVQA (Trustworthy MLLM):** Led the construction of a benchmark for "evidence attribution" in MLLMs. Built an automated pipeline covering 711 long documents and 2,000 fine-grained QA pairs. Successfully quantified the "attribution decoupling" (correct answer but wrong evidence) hallucination in SOTA

models.

- **AgenticOCR (Efficient RAG):** Optimized document parsing for multimodal RAG. Implemented a query-driven dynamic extraction paradigm to replace static OCR, mitigating attention overload in long-context retrieval and enhancing token efficiency.

## Huawei Technologies Co., Ltd.

Aug 2025 – Dec 2025

Intern — Advisor: Pengfei Li

- **Ascend C Operator Generation:** Developed an automated "Generation-Error-Correction" loop for NPU operators. Synthesized 10k+ high-quality CoT and alignment data for SFT/DPO. Improved compilation pass rate by 24% and accuracy by 18% in agent-assisted scenarios (**Work accepted by ACL 2026 Findings**).
- **Linear Attention and On-device Optimization:** Linearized Qwen2.5-1.5B by replacing full-attention with GDN (Gated Delta Net). Optimized Beam Search-based layer searching for sub-network discovery, providing a viable solution for LLM deployment on resource-constrained devices.

## Research Experience

### Efficient Edge-Cloud Collaborative Inference for LLMs

Mar 2025 – Dec 2025

Research Fields: Reasoning Planning, Multi-Agent Systems

Advisor: Prof. Wanyu Lin

- **Algorithm Implementation:** Proposed a resource-adaptive edge-cloud collaborative inference framework. Fine-tuned a lightweight edge planner to decompose complex queries into fine-grained, DAG-based subtasks. Designed a dynamic routing mechanism to realize dependency-aware collaborative reasoning by executing simple tasks in parallel on the edge while offloading high-difficulty tasks to the cloud. (**Accepted by ICML 2026**).

### Audio-Driven 3D Head Video Generation

Dec 2023 – Jun 2024

Research Fields: Computer Vision, 3D Reconstruction, GAN

Advisor: Prof. Yan Xing

- **Algorithm Implementation:** Integrated GMM-HMM models for voice command recognition and utilized 2D feature landmarks to drive 3D facial reconstruction. Engineered a GAN-based network architecture to achieve precise control over the generation of subtle facial expressions at the Action Unit (AU) level. The methodology successfully led to **1 National Invention Patent** and was awarded the **Outstanding Undergraduate Thesis**.

## Project Experience

### RAG-Based Knowledge Question-Answering System [GitHub](#)

Mar 2025 – Apr 2025

- Designed a configurable LLM-based QA system with a dynamic routing mechanism to seamlessly switch among multiple retrieval strategies based on specific application scenarios.
- Introduced a multi-agent collaboration mechanism decoupling key information extraction from final answer generation, significantly improving retrieval recall and QA accuracy for vertical-domain documents.

### LLM-Driven Heuristic Auxiliary Learning System [Streamlit App](#)

Nov 2024 – Dec 2024

- Developed an intelligent tutoring platform utilizing Streamlit. Employed heuristic step-by-step guidance (multiple-choice prompts and logical hints) instead of direct answer generation to foster users' independent problem-solving capabilities.
- Integrated a prerequisite knowledge graph review module and a gamified card-collection mechanism to systematically enhance user engagement and learning retention.

## Honors and Awards

### Outstanding Undergraduate Thesis Award

Jun 2024

Top 2% students in Department of Mathematics, Hefei University of Technology

### Outstanding Student Scholarship

Dec 2023

Top 10% students in Department of Mathematics, Hefei University of Technology

### Outstanding Student Scholarship for Excellence in Scientific & Technological Activities

Dec 2023

|                                                                                                                         |                  |
|-------------------------------------------------------------------------------------------------------------------------|------------------|
| Top 2% students in Department of Mathematics, Hefei University of Technology                                            |                  |
| <b>Silver Award of the 8th Anhui Province “Internet Plus” Innovation and Entrepreneurship Competition</b>               | <i>Nov 2022</i>  |
| Department of Education of Anhui Province                                                                               |                  |
| <b>Gold Award of the 8th Hefei University of Technology “Internet Plus” Innovation and Entrepreneurship Competition</b> | <i>July 2022</i> |
| Hefei University of Technology                                                                                          |                  |
| <b>Outstanding Student Scholarship</b>                                                                                  | <i>Dec 2021</i>  |
| Top 10% students in Department of Mathematics, Hefei University of Technology                                           |                  |
| <b>Merit Student</b>                                                                                                    | <i>Dec 2021</i>  |
| Top 2% students in Department of Mathematics, Hefei University of Technology                                            |                  |
| <b>Outstanding Young Volunteer</b>                                                                                      | <i>Dec 2021</i>  |
| Hefei University of Technology                                                                                          |                  |
| <b>Outstanding Volunteer</b>                                                                                            | <i>Oct 2021</i>  |
| The Yangtze Crossing Campaign Memorial                                                                                  |                  |

## Technical Skills

---

**Languages:** Python, SQL (Oracle & SQL Server), Latex, JavaScript, HTML/CSS, C/C++, MATLAB, R

**Developer Tools:** VS Code, PyCharm, Git, Linux, IntelliJ IDEA

**Libraries/Frameworks:** PyTorch, TensorFlow, OpenCV, NumPy, Pandas, Vue.js

## Languages

---

|                                                                               |                          |
|-------------------------------------------------------------------------------|--------------------------|
| <b>Mandarin</b>                                                               | <i>Native speaker</i>    |
| Certificate of Putonghua Test, Band 2, Level 1, State Language Work Committee |                          |
| <b>English</b>                                                                | <i>Highly proficient</i> |
| IELTS: 6.5                                                                    |                          |

## Extra-curricular Activities

---

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                            |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|
| <b>Volunteer Activities</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | <i>Sep 2020 — Jun 2024</i> |
| <ul style="list-style-type: none"> <li>◦ Campus anti-epidemic volunteer<br/>Responsible for delivering food, drinking water, and daily necessities for students.</li> <li>◦ Volunteer teacher<br/>Provided teaching at Hefei Shushan Primary School and in communities in Binhu New Area</li> <li>◦ Volunteer service in Anhui Provincial Library<br/>Shelved children’s books; Provided children storytelling picture book lessons to expand their imagination.</li> <li>◦ The Yangtze Crossing Campaign Memorial service volunteer</li> </ul> |                            |
| <b>Minister of National Folk Art Association</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | <i>Sep 2021 — Jun 2022</i> |
| Hefei University of Technology                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                            |
| <b>Head of the publicity department of Zhaoyang Volunteer Service Association, Student Union</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                | <i>Sep 2021 — Jun 2022</i> |
| Department of Mathematics, Hefei University of Technology                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                            |
| <b>Member of Hefei University of Technology Chinese orchestra</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | <i>Sep 2021 — Jun 2024</i> |
| The principal first Erhu player, Hefei University of Technology                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                            |