

JUNBO NIU

📞 18810092213 ✉ 21376334@buaa.edu.cn 📄 Personal Site 🌐 github.com/NiuJunbo2002

Education

Peking University

Ph.D. Candidate, Electrical Engineering and Computer Science

Beijing, China

Starting 2025.09

- Major in Artificial Intelligence.
- Advised by **Prof. Wentao Zhang**.

Beihang University

B.S., Automation Science

Beijing, China

2021.09 – 2025.06

- Major in Computer Vision and Pattern Recognition.
- GPA: 3.85/4.00, Rank: 1/203 (Top 1%)

Selected Publications

OVOBench: How Far is Your Video-LLMs from Real-World Online Video Understanding? (CVPR 2025). 📄

- Explored how to bridging the gap between model performance and human-level online video understanding in video-based AI.
- Video-LLMs: Benchmarking and Streaming/Online Video Understanding.
- **Junbo Niu***, Yifei Li*, Ziyang Miao, Chunjiang Ge, Jiaqi Wang[†] *et al.*

Native Visual Understanding: Resolving Resolution Dilemmas in Vision-Language Models (NeurIPS 2025). 📄

- Identified critical resolution dilemmas in the domain of VLMs.
- Proposed and open-sourced NativeRes-LLaVA, a framework enabling robust native resolution visual encoding by preserving crucial image details, demonstrating superiority over existing methods and advancing open-source capabilities in this domain.
- **Junbo Niu***, Yuanhong Zheng*, Ziyang Miao, Hejun Dong, Wentao Zhang[†] *et al.*

InternLM-XComposer2.5-OmniLive (Technical Report) 📄

- The multimodal real-time interactive model(InternLM-XComposer 2.5-OmniLive). This model can observe and understand the external world in real-time through visual and auditory inputs, automatically form long-term memory of the observed content, and engage in conversational interactions with human users via voice.
- A Comprehensive Multimodal System for **Long-term Streaming Video** and **Audio Interactions**.

Projects

MinerU 2.0 🌐

2024.11 – 2025.07

- **36.9K GitHub stars**, showcasing significant community impact and adoption.
- Engineered a **unified, end-to-end model for general OCR tasks**, significantly boosting performance and operational efficiency.
- Designed and implemented the **core model training framework**, actively exploring and integrating novel model architectures.
- Streamlined data pipelines and **led the optimization of the model architecture's training process**.
- **Core Contributor**.

ShareGPT4Video 🌐

2024.06 – 2024.09

- Pioneered the DataCoT method for generating high-quality visual instruction data, achieving substantial cost reductions by leveraging open-source large language models.

Honours and Awards

Outstanding student of Beihang University

09/2022

Merit Student of Beihang University

12/2022

Star of “Yu-Yuan”, Top 2 of 600+ ,The highest honor of the department.

12/2024

National Scholarship, (Ministry of Education, People's Republic of China)

12/2024

Experience

| | |
|---|--------------------------|
| LEarning And Perception Laboratory (LEAP Lab) , Tsinghua University <i>Research Intern, Supervised by Prof. Gao Huang and Ph.D. Chunjiang Ge</i> <ul style="list-style-type: none">• MLLM | 2023.09 – 2024.05 |
| Large Model Center, Shanghai AI Lab <i>Research Intern, Supervised by Dr. Jiaqi Wang.</i> <ul style="list-style-type: none">• Video-LLMs• Online Video Understanding | 2024.06 – 2024.10 |
| OpenDataLab, Shanghai AI Lab <i>Research Intern, Supervised by Dr. Conghui He.</i> <ul style="list-style-type: none">• OCR-Free Visual Understanding• OCR-Free Model Pretrain• Visual Feature Compressing | 2024.10 – 2025.07 |

Skills

- Languages:** Familiar with Python and C++/C
- Tools:**Pytorch, Git, Linux, OpenCV, Unity3D
- Maths:** Familiar with mathematics analysis, linear algebra, probability theory, complex function and integral transform and discrete mathematics
- Language:**English - CET-6: 503/710, Mandarin - Native speaker
- Miscellaneous:** Basketball (PB: Departmental champion), Outdoor sport