

JUNBO NIU

📞 18810092213 ✉ 21376334@buaa.edu.cn 📄 Personal Site 🌐 github.com/Niujunbo2002

Education

Beihang University

BS, Automation Science

Beijing, China

2021 – Expected 2025

- GPA: 3.85/4.00, Rank: 1/203 (Top 1%)
- Will pursue a Ph.D. in Artificial Intelligence at PKU (Advised by **Prof. Wentao Zhang** and **Prof. Bin Cui**)

Publications

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

- Explore 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.*

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**
- Pan Zhang, ... , **Junbo Niu**, Jiaqi Wang[†] *et al.*

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 09/2023 – 03/2024

Research Intern, Supervised by Prof. Gao Huang and Ph.D. Chunjiang Ge

- MLLM

Large Model Center, Shanghai AI Lab 06/2024 – 12/2024

Research Intern, Supervised by Dr. Jiaqi Wang.

- Video-LLMs
- Online Video Understanding

OpenDataLab, Shanghai AI Lab 12/2024 – 2/2025

Research Intern, Supervised by Dr. Conghui He.

- OCR-Free Visual Understanding
- OCR-Free Model Pretrain
- Visual Feature Compressing

Projects

Multimodal Instruction Data Generation (09/2023 - 02/2024)

- Developed the DataCoT method to generate high-quality visual instruction data at significantly reduced costs using open-source large language models.

ShareGPT4Omni Code Pipeline Refine (08/2024 - 12/2024)

- Build a more general video/image captioning framework.

MLLM Pretrain for OCR(GOT-OCR2.0) (12/2024 - 2/2025)

- Explore the use of a Unified End-to-end Model for general OCR tasks.
- Build model training framework and explore model architecture.

Skills

Languages: Familiar with Python and C++/C, Shell, R

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