

# Boyuan Sun

Phone: +86-15253758970 Email: [sbysbysby123@gmail.com](mailto:sbysbysby123@gmail.com)

Personal Website: [bbbchan.github.io](http://bbbchan.github.io) Personal Blog: [bbbchan.com](http://bbbchan.com)

---

## Profile

Boyuan Sun is a third-year Ph.D. candidate at the Media Computing Laboratory, Nankai University, advised by [Prof. Qibin Hou](#) and [Prof. Ming-Ming Cheng](#). He received his B.S. degree from Xidian University in 2021 and is currently pursuing a joint Master-Ph.D. program at the College of Computer Science, Nankai University. His research has been cited 500+ times on [Google Scholar](#), and his [GitHub](#) repositories have accumulated 1000+ stars. Research interests include computer vision and deep learning, with a particular focus on multimodal large language models, video understanding, and agentic model training.

---

## Education & Experience

- **ByteDance - Douyin**  
Research Intern – Streaming Video Understanding Beijing, China  
Apr. 2026 – Present
  - **Shanghai AI Laboratory**  
Research Intern – MLE Agent Workflow Shanghai, China  
Dec. 2025 – Apr. 2026
  - **Alibaba - Tongyi Lab**  
Research Intern – Token Compression, Video Understanding Hangzhou, China  
Sep. 2024 – Dec. 2025
  - **Media Computing Lab (VCIP), Nankai Univ.**  
Computer Science, Joint Master-Ph.D. Tianjin, China  
Sep. 2021 – Present
  - **Xidian University**  
Software Engineering, B.S. Xi'an, China  
Sep. 2017 – Jul. 2021
- 

## Publications (First/Co-First)

- [1] **Boyuan Sun**, Yuqi Yang, Le Zhang, Ming-Ming Cheng, Qibin Hou. "CorrMatch: Label Propagation via Correlation Matching for Semi-Supervised Semantic Segmentation." **CVPR 2024**.
  - [2] Jiaxing Zhao\*, **Boyuan Sun\***, Xiang Chen, Xihan Wei. "Facial Dynamics in Video: Instruction Tuning for Improved Facial Expression Perception and Contextual Awareness." **AAAI 2026**.
  - [3] **Boyuan Sun**, Bowen Yin, Yuanming Li, Xihan Wei, Qibin Hou. "See What I Mean: Aligning Vision and Language Representations for Video Fine-grained Object Understanding." **CVPR 2026**.
- 

## Submitted Works (First/Co-First)

- [1] **Boyuan Sun**, Modi Jin, Bowen Yin, Qibin Hou. "Depth Anything at Any Condition." **Submitted to TPAMI**.
  - [2] **Boyuan Sun**, Jiaxing Zhao, Xihan Wei, Qibin Hou. "LLaVA-Scissor: Token Compression with Semantic Connected Components for Video LLMs." **Submitted to TMM**.
  - [3] **Boyuan Sun\***, Jiaxing Zhao\*, Xiang Chen, Xihan Wei, Qibin Hou. "LLaVA-Octopus: Unlocking Instruction-Driven Adaptive Projector Fusion for Video Understanding."
- 

## Technical Reports and Open-source Projects

- [1] HumanOmni: A Large Vision-Speech Language Model for Human-Centric Video Understanding
  - [2] HumanOmniV2: From Understanding to Omni-Modal Reasoning with Context
  - [3] cc-mini: Ultra-light Harness scaffolding for AI agents. <https://github.com/e10nMa2k/cc-mini>
- 

## Honors & Awards

- |                                       |                     |
|---------------------------------------|---------------------|
| ACM/ICPC National Invitational        | <b>Bronze Medal</b> |
| ACM/ICPC Sichuan Provincial           | <b>Silver Medal</b> |
| ACM/ICPC Shaanxi Provincial           | <b>Bronze Medal</b> |
| First-Class Scholarship, Xidian Univ. | 2018                |
| National Scholarship, China           | 2019                |
| First-Class Scholarship, Nankai Univ. | 2024                |
-