Kunchang Li

kc.li@siat.ac.cn · (+86) 15652583072 · Google · GitHub

Education

University of Chinese Academy of Sciences, Shenzhen, China 2020 – Present

Ph.D. student in Computer Application Technology at Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, advised by Prof. Yu Qiao & Prof. Yali Wang.

- Overall GPA: 3.73/4.0 Rank: Not provided

Beihang University, Beijing, China 2016 – 2020

B.E. in Software Engineering at the School of Software.

- Overall GPA: 3.85/4.0 Rank: 4/165

Honors

- **Top 10 Outstanding** Graduate Students in SIAT, UCAS 2023
- **National award scholarship** 2023
- **Dean Scholarship** of University of Chinese Academy of Sciences 2022, 2023
- **Merit Student** at University of Chinese Academy of Sciences 2021, 2022
- **1st** Place in Forecasting Challenge (ECCV2022 Ego4D Workshop) 2022
- **2nd** Place in Action Recognition in the Dark Challenge (CVPR2022 UG2+ Workshop) 2022
- **1st** Place in Semantic Segmentation of Remote Sensing Images (CCF BDCI Contest) 2021
- **Comprehensive Grand Prize** of CCF BDCI Contest 2021
- **Excellent** Higher Education Graduate of Beijing Municipality 2020
- Grand Prize of Social Work Scholarship, Grand Prize of Study Excellence Scholarship, Merit Student and Honor student at Beihang University 2017, 2018, 2019

Research Interests

- Vision Backbone (Efficient Architecture Design, Large-scale Pre-training)
- Video Understanding (Action Recognition)
- Multimodal Learning & Large Language Model
- Video Generation (Minute-level Vlog)

Internships

Shanghai AI Lab, Shanghai, China Nov. 2021 – Present

Advisors: Yali Wang, Limin Wang and Yi Wang
General Video Foundation Model; Large-scale Pre-training; Multimodal Learning

SenseTime, Beijing, China Feb. 2021 – Nov. 2021

Advisors: Guanglu Song and Yu Liu
Efficient Architecture Design; Video Understanding

MEGVII, Beijing, China Oct. 2019 – Jan. 2020

Acceleration for Face Recognition Model; Model Reproduction for MegEngine

Selected Papers

See the full paper list here. * refers to the **co-first authors**. All the code can be found on GitHub.

[1] **Kunchang Li**, Xinhao Li, et al., "VdeoMamba: State Space Model for Efficient Video Understanding." Arxiv:2403.06977.

[2] **Kunchang Li**, Yali Wang, et al., "MVBench: A Comprehensive Multi-modal Video Understanding Benchmark." **CVPR2024**.

[3] **Kunchang Li***, Yali Wang*, et al., "UniFormer: Unifying convolution and self-attention for visual recognition." **TPAMI2023**.
[4] **Kunchang Li**, Yali Wang, et al., "Unmasked Teacher: Towards Training-Efficient Video Foundation Models." **ICCV2023 (Oral Presentation, Top 2%)**.
[5] **Kunchang Li**, Yali Wang, et al., "UniFormerV2: Unlocking the Potential of Image ViTs for Video Understanding." **ICCV2023**.
[6] **Kunchang Li***, Yali Wang*, et al., "UniFormer: Unified Transformer for Efficient Spatiotemporal Representation Learning." **ICLR2022 (Review score 8868, Top 3%)**.
[7] **Kunchang Li***, Yali Wang*, et al., "CT-Net: Channel tensorization network for video classification." **ICLR2021**.
[8] **Kunchang Li***, Yinan He*, et al., "VideoChat: Chat-Centric Video Understanding." Arxiv:2305.06355.
[9] Yi Wang*, **Kunchang Li***, et al., "InternVideo: General Video Foundation Models via Generative and Discriminative Learning." Arxiv:2212.03191.
[10] Yi Wang*, **Kunchang Li***, et al., "InternVideo2: Scaling Foundation Models for Multimodal Video Understanding." Arxiv:2403.15377.
[11] Junhao Zhang*, **Kunchang Li***, et al., "MorphMLP: A Self-Attention Free, MLP-Like Backbone for Image and Video." **ECCV2022**.
[12] Zhuofan Zong*, **Kunchang Li***, et al., "Self-slimmed vision transformer." **ECCV2022**.
[13] Xiaoxiao Sheng*, **Kunchang Li***, et al., "A Progressive Difference Method for Capturing Visual Tempos on Action Recognition." **TCSVT2022**.
[14] Shaobin Zhuang, **Kunchang Li**, et al., "Vlogger: Make Your Dream A Vlog." **CVPR2024**.
[15] Yi Wang*, Yinan He*, Yizhuo Li*, **Kunchang Li**, et al. "InternVid: A Large-scale Video-Text Dataset for Multimodal Understanding and Generation." **ICLR2024**.
[16] Ziteng Cui, **Kunchang Li**, et al., "Illumination Adaptive Transformer." **BMVC2022**.
[17] Renrui Zhang, Ziyu Guo, Wei Zhang, **Kunchang Li**, et al., "PointCLIP: Point Cloud Understanding by CLIP." **CVPR2022**.
[18] Renrui Zhang, Rongyao Fang, Peng Gao, Wei Zhang, **Kunchang Li**, et al., "Tip-adapter: Training-free clip-adapter for better vision-language modeling." **ECCV2022**.
[19] Guo Chen, Sen Xing, Zhe Chen, Yi Wang, **Kunchang Li**, et al., "InternVideo-Ego4D: A Pack of Champion Solutions to Ego4D Challenges." **ECCVW2022**.
[20] Weicong Su, Yali Wang, **Kunchang Li**, et al., "Hybrid token transformer for deep face recognition" **PR2023**.

PROJECTS

- **Grounded-Segment-Anything**: GitHub Trending with **13k+** stars.
- **Ask-Anything**: GitHub Trending with **2k+** stars (**150+** citations).
- **InternGPT**: GitHub Trending with about **3k** stars.
- **UniFormer**: Efficient vision backbone (**350+** citations).
- **UniFormerV2**: Plug-and-play temporal module with strong performances (**60+** citations).
- **Unmasked Teacher**: Efficient and powerful video foundation models (**40+** citations).
- **InternVideo**: General video foundation models (**130+** citations).
- **IAT**: 90K parameters to adapt light (**80+** citations).
- **Seg4Fun**: **Top-1** solution for CCF BDCI Segmentation.

SERVICES

Conference Reviewer: ICLR2023/2024, CVPR2023/2024, ICCV2023, NeurIPS2023, ICML2024, ECCV2024
Journal Reviewer: TPAMI, IJCV, PR, NN, JVCI
Talk: AI Drive 2022, AI Time 2022, AI Time 2023, others (FDU, UCAS, KAUST...)