Haodong Duan

✉ dhd.efz@gmail.com · ⌂ HomePage · 🏛 Google Scholar

Professional Experience

Post-Doctoral Researcher, Shanghai AI Laboratory   **2023 – Present**

Work on evaluating large language models and large multi-modal models, supervised by Prof. Dahua Lin, mentored by Dr. Kai Chen.

Applied Scientist Intern, AWS AI   **2022**

Work on skeleton action recognition in the wild, mentored by Dr. Mingze Xu and Dr. Alessandro Bergamo.

Research Intern, Shanghai AI Laboratory   **2021 – 2022**

Develop and maintain the OpenSource codebase MMAction2, mentored by Dr. Kai Chen.

Research Intern, Sensetime   **2017 – 2019**

Work on human pose estimation (skeleton keypoints and contour keypoints), mentored by Dr. Wentao Liu.

Education

Chinese University of HongKong, HongKong   **2019 – 2023**

Working on *Video Understanding*, supervised by Dahua Lin
Ph.D. in Information Engineering

Peking University, Beijing   **2015 – 2019**

GPA 3.77/4.00, rank 1st in Data Science students
Undergraduate in Data Science, Yuanpei College

Projects

OpenSource Codebases:
The main contributor and maintainer of MMAction2, PYSKL, OpenCompass, VLMEvalKit.

- MMAction2: An open-source codebase for video understanding and action recognition. (★ 4.2K, ⑃ 1.2K)
- PYSKL: An open-source codebase for skeleton-based action recognition. (★ 950, ⑃ 181)
- OpenCompass: An open-source codebase for evaluating large language models. (★ 3.8K, ⑃ 406)
- VLMEvalKit: An open-source codebase for evaluating multi-modal models (★ 1.1K, ⑃ 157). We also host the OpenVLM Leaderboard (a huggingface space, ♥ 422) based on this toolkit.

LLM / VLM Projects:
- A main contributor of InternLM series: Mainly works on Subjective Evaluation, Long-Context Evalution and Multi-Turn Dialogue Evaluation.
- A main contributor of InternLM-XComposer series: Mainly works on the evalution of InternLM-XComposer.

Preprints

Xiangyu Zhao, Xiangtai Li, **Haodong Duan**, Haian Huang, Yining Li, Kai Chen, Hua Yang
MG-LLaVA: Towards Multi-Granularity Visual Instruction Tuning

Haodong Duan, Jiaqi Wang, Kai Chen, Dahua Lin
DG-STGCN: Dynamic Spatial-Temporal Modeling for Skeleton-based Action Recognition

PUBLICATIONS

Xinyu Fang, Kangrui Mao, **Haodong Duan**[†], Xiangyu Zhao, Yining Li, Dahua Lin, Kai Chen
MMBench-Video: A Long-Form Multi-Shot Benchmark for Holistic Video Understanding
(NeurIPS 2024 Dataset)

Yuxuan Qiao, **Haodong Duan**[†], Xinyu Fang, Junming Yang, Lin Chen, Songyang Zhang, Jiaqi Wang, Dahua Lin, Kai Chen
Prism: A Framework for Decoupling and Assessing the Capabilities of VLMs (NeurIPS 2024)

Xiaoyi Dong, Pan Zhang, Yuhang Zang, Yuhang Cao, Bin Wang, Linke Ouyang, Songyang Zhang, **Haodong Duan**, Wenwei Zhang, Yining Li, Hang Yan, Yang Gao, Zhe Chen, Xinyue Zhang, Wei Li, Jingwen Li, Wenhai Wang, Kai Chen, Conghui He, Xingcheng Zhang, Jifeng Dai, Yu Qiao, Dahua Lin, Jiaqi Wang
InternLM-XComposer2-4KHD: A Pioneering Large Vision-Language Model Handling Resolutions from 336 Pixels to 4K HD (NeurIPS 2024)

Pengcheng Chen, Jin Ye, Guoan Wang, Yanjun Li, Zhongying Deng, Wei Li, Tianbin Li, **Haodong Duan**, Ziyan Huang, Yanzhou Su, Benyou Wang, Shaoting Zhang, Bin Fu, Jianfei Cai, Bohan Zhuang, Eric J Seibel, Junjun He, Yu Qiao
GMAI-MMBench: A Comprehensive Multimodal Evaluation Benchmark Towards General Medical AI
(NeurIPS 2024 Dataset)

Lin Chen, Xilin Wei, Jinsong Li, Xiaoyi Dong, Pan Zhang, Yuhang Zang, Zehui Chen, **Haodong Duan**, Bin Lin, Zhenyu Tang, Li Yuan, Yu Qiao, Dahua Lin, Feng Zhao, Jiaqi Wang
Sharegpt4video: Improving video understanding and generation with better captions (NeurIPS 2024 Dataset)

Lin Chen, Jinsong Li, Xiaoyi Dong, Pan Zhang, Yuhang Zang, Zehui Chen, **Haodong Duan**, Jiaqi Wang, Yu Qiao, Dahua Lin, Feng Zhao
Are We on the Right Way for Evaluating Large Vision-Language Models? (NeurIPS 2024)

Haodong Duan, Junming Yang, Yuxuan Qiao, Xinyu Fang, Lin Chen, Yuan Liu, Xiaoyi Dong, Yuhang Zang, Pan Zhang, Jiaqi Wang, Dahua Lin, Kai Chen
VLMEvalKit: An Open-Source Toolkit for Evaluating Large Multi-Modality Models (MM 2024)

Yuan Liu, **Haodong Duan**[‡], Yuanhan Zhang, Bo Li, Songyang Zhang, Wangbo Zhao, Yike Yuan, Jiaqi Wang, Conghui He, Ziwei Liu, Kai Chen, Dahua Lin
Mmbench: Is your multi-modal model an all-around player? (ECCV 2024 Oral)

Hongwei Liu, Zilong Zheng, Yuxuan Qiao, **Haodong Duan**, Zhiwei Fei, Fengzhe Zhou, Wenwei Zhang, Songyang Zhang, Dahua Lin, Kai Chen
MathBench: Evaluating the Theory and Application Proficiency of LLMs with a Hierarchical Mathematics Benchmark (ACL 2024 Findings)

Chonghua Wang, **Haodong Duan**[†], Songyang Zhang, Dahua Lin, Kai Chen
Ada-LEval: Evaluating long-context LLMs with length-adaptable benchmarks (NAACL 2024)

Haodong Duan[†], Jueqi Wei, Chonghua Wang, Hongwei Liu, Yixiao Fang, Songyang Zhang, Dahua Lin, Kai Chen
BotChat: Evaluating LLMs' Capabilities of Having Multi-Turn Dialogues (NAACL 2024 Findings)

Junting Pan, Keqiang Sun, Yuying Ge, Hao Li, **Haodong Duan**, Xiaoshi Wu, Renrui Zhang, Aojun Zhou,

Zipeng Qin, Yi Wang, Jifeng Dai, Yu Qiao, Hongsheng Li
Journeydb: A benchmark for generative image understanding (NeurIPS 2023 Dataset)

Haodong Duan, Mingze Xu, Bing Shuai, Davide Modolo, Zhuowen Tu, Joseph Tighe, Alessandro Bergamo
SkeleTR: Towards Skeleton-based Action Recognition in the Wild (ICCV 2023)

Yujie Zhou, **Haodong Duan**, Anyi Rao, Bing Su, Jiaqi Wang
Self-supervised Action Representation Learning from Partial Spatio-Temporal Skeleton Sequences (AAAI 2023)

Haodong Duan, Yue Zhao, Kai Chen, Yuanjun Xiong, Dahua Lin
Mitigating Representation Bias in Action Recognition: Algorithms and Benchmarks (ECCVW 2022)

Haodong Duan, Jiaqi Wang, Kai Chen, Dahua Lin
PYSKL: Towards Good Practices for Skeleton Action Recognition (MM 2022)

Haodong Duan, Yue Zhao, Kai Chen, Dahua Lin, Bo Dai
Revisiting Skeleton-based Action Recognition (CVPR 2022 Oral)

Haodong Duan, NanXuan Zhao, Kai Chen, Dahua Lin
TransRank: Self-supervised Video Representation Learning via Ranking-based Transformation Recognition (CVPR 2022 Oral)

Jintao Lin, **Haodong Duan**, Kai Chen, Dahua Lin, Limin Wang
OCSampler: Compressing Videos to One Clip with Single-step Sampling (CVPR 2022)

Haodong Duan, Yue Zhao, Yuanjun Xiong, Wentao Liu, Dahua Lin
Omni-sourced Webly-supervised Learning for Video Recognition (ECCV 2020)

Haodong Duan, Kwanyee Lin, Sheng Jin, Wentao Liu, Chen Qian, Wanli Ouyang
TRB: A Novel Triplet Representation for Understanding 2D Human Body (ICCV 2019 Oral)

[†] Corresponding Author, [‡] Project Lead

Professional Services

- **Conference Reviewer:** ICCV[21-23], AAAI[22-25], CVPR[22-24], ECCV[22-24], NeurIPS[22-24], WACV23, ICML23, ICLR[24-25], EuroGraphics23
- **Journal Reviewer:** TPAMI, IJCV, TIP, PR, TMM, TOMM, TCSVT, Sensors, SPL, JVCIR

Grants

- **Shanghai Postdoctoral Excellence Program**, No. 2023023
- **China Postdoctoral Science Fund**, No. 2024M751559

Language Skills

- TOEFL iBT test: 104pt (Reading: 30, Listening: 28, Speaking: 20, Writing: 26)
- GRE test: 322pt (Verbal: 152, Quantitative: 170)