Haodong Duan

✉ dhd.efz@gmail.com · ☎ +86 18211165536 · ⌂ HomePage

Education

Peking University, Beijing   **2015 – 2019**

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

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

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

Research Interests

My research interests lie in the area of computer vision and video understanding. In particular, I'm focusing on efficient video understanding (data-efficient & computationally efficient), based on skeleton action recognition.

Research Projects

Unified Framework for Skeleton Action Recognition in the Wild   **2022**

Design a unified framework **SkeleTR** to handle instance-level, group-level, video-level skeleton action recognition in the wild, which outperforms the GCN SOTAs by a large margin on all benchmarked tasks.

Skeleton Action Recognition with Dynamic Group-wise GCN, Paper   **2021 – 2022**

Design a dynamic GCN for skeleton action recognition, that requires no pre-defined skeleton topology. It achieves strong recognition performance, surpasses previous SOTAs on NTURGB+D and K400.

Video Self-supervised Learning via Ranking-based Transformation Recognition, Paper   **2021**

Show the great potential of RecogTrans (recognizing transformations applied to video clips) video SSL by introducing a unified Ranking-based formulation. The proposed method significantly outperforms previous RecogTrans approaches on action recognition (UCF Top1 +6%) and video retrieval (UCF R@1 +20%).

Efficient Video Recognition for Untrimmed Videos, Paper, Code   **2021**

Propose OCSampler for untrimmed video recognition. It samples frames from candidates to form one representative clip. The framework (w. R50) achieves 82.2% Top-1 with 1-clip testing (only 52 GFLOPs / video).

Skeleton Action Recognition with 3D ConvNets, Paper, Dataset, Code   **2020 – 2021**

Devise a novel 3D-ConvNet based paradigm (**PoseC3D**: 2D keypoint heatmaps → 3D heatmap volumes → 3D-CNN recognizer) for skeleton action recognition. PoseC3D outperforms previous skeleton action recognition approaches by a considerable margin across various benchmarks (NTURGB+D, Kinetics, *etc.*).

Mitigating Unwanted Bias in Action Recognition, Paper, Dataset   **2020**

Demonstrate that deep learning based video recognition is biased towards factors like scene / objects. Create a quantitative benchmark to evaluate the bias, mitigate the bias with adversarial training and diversified web data.

Omni-sourced Webly-supervised Video Recognition, Paper, Dataset, Code   **2019 – 2020**

Propose **OmniSource** for webly supervised video recognition, which utilizes various kinds of web data, including images, trimmed / untrimmed videos for trimmed video recognition. Achieve 83.6% Top-1 on Kinetics400.

Triplet Representation for Human Body, Paper, Dataset   **2018 – 2019**

Design a triplet representation named TRB (and its estimation method) to represent 2D human body, which includes both human pose and shape information. The representation can be used in human shape editing.

Experience

Research Intern, Sensetime   **2017 – 2019**

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

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

Develop and maintain the OpenSource codebase MMAction2, 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.

Preprints

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

Publications

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)

OpenSource Projects

The main contributor and maintainer of MMAction, MMAction2, and PYSKL.

Professional Services

Conference Reviewer: ICCV21, AAAI[22-23], CVPR[22-23], ECCV22, NeurIPS22, WACV23.
Journal Reviewer: TCSVT, SPL, JVCIR.

Language Skills

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