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 Projects

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

Design a triplet representation named TRB (as well as 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.

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

Propose a framework (**OmniSource**) for webly supervised video recognition, which can utilize various kinds of web medias, including images, trimmed videos, untrimmed videos for trimmed video recognition. Achieve 83.6% Top-1 Accuracy on Kinetics400 with SOTA algorithm and our framework.

Mitigating Unwanted Bias in Action Recognition 2020

Demonstrate that the deep learning based video recognition models are biased towards factors like scene or objects. Establish a new quantitative benchmark to evaluate such bias and propose to mitigate such bias with adversarial training and diversified web data.

Skeleton-based Action Recognition with 3D-CNN, Paper, Dataset, Code 2020 – 2021

Devise a new pipeline (**PoseC3D**) for skeleton-based action recognition, using Top-Down pose estimators as pose extractors, 3D heatmap volumes as the representation, and 3D-CNN as the recognizer. PoseC3D outperforms previous solutions for skeleton-based action recognition by a considerable margin.

Efficient Video Recognition for Untrimmed Videos 2021

Propose an efficient framework for untrimmed video recognition, which samples frames from frame candidates to form one representative clip. The framework (w. R50 backbone) can achieve 82.2% Top-1 Accuracy with 1-clip testing (the computational cost is only 52 GFLOPs / video).

OpenSource Projects

The maintainer of MMAction and MMAction2.

Professional Activities

Conference Reviewer for ICCV2021.

Language Skills

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