بسم الله الرحمن الرحیم

شرکت مهندسی نرم افزاری هلو

گزارش مربوط به روش­های مجموعه های داده­ای مربوط به Action Recognition

کاری از امیرعلی نسیمی

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| وفق پذیری | دقت گزارش شده | توضیح اضافه | ساختار شبکه | لینک | منبع | نام روش | مجموعه داده | شماره |
| مناسب | - | نیاز به دو مجموعه داده MPII و [Ntu -RGB](http://rose1.ntu.edu.sg/datasets/actionrecognition.asp) جهت آموزش | Conv | [link](https://github.com/KanghyuLee/Deep-human-action-recognition/tree/master) | GitHub | DHCR | [MPII](http://human-pose.mpi-inf.mpg.de/) | 1 |
| بسیار مناسب | [link](https://github.com/Fanbenchao/Action-Recognition-based-on-pose-estimation) | Action Recognition based on pose estimation | 2 |
| مناسب | [link](https://www.youtube.com/watch?v=MNEZACbFA4Y&t=6s) | نیاز به دو مجموعه داده MPII و Human3.6Mجهت آموزش | [Conv](https://arxiv.org/pdf/1912.08077.pdf) | [link](https://github.com/dluvizon/deephar) | Deephar | 3 |
| متوسط | [link](http://chengao.vision/SDN/) | [link](http://papers.nips.cc/paper/8372-why-cant-i-dance-in-the-mall-learning-to-mitigate-scene-bias-in-action-recognition.pdf) | [Conv](http://chengao.vision/SDN/images/overview.png) | [link](http://chengao.vision/SDN/) | SDN | 4 |
| [link](https://github.com/vt-vl-lab/SDN) |
| مناسب | 37.8 | - | Inception | [link](https://github.com/rohitgirdhar/AttentionalPoolingAction) | Attentional Pooling for Action Recognition | 5 |
| [link](https://rohitgirdhar.github.io/AttentionalPoolingAction/) | Article Page |
| بسیار مناسب – نیاز به آموزش بر مجموعه داده مورد نظر | [link](https://rohitgirdhar.github.io/ActionTransformer/assets/suppl/pred.pdf) | ورودی بصورت فریم­های زمانی از ویدئو بوده که به شبکه R-CNN متصل می­شود. سپس ساختار خود را بصورت شبکه­های دو شاخه جلو می­برد. توضیح بیشتر در فایل ویدئو: [link](https://rohitgirdhar.github.io/ActionTransformer/assets/video_present.mp4) | Fast R-CNN - Transformer | [link](https://rohitgirdhar.github.io/ActionTransformer/) | VATN | [AVA](http://research.google.com/ava/) | 6 |
| تنها Back Bone بوده و یادگیری انتقالی ندارد | [link](https://github.com/ppriyank/Video-Action-Transformer-Network-Pytorch-) | GitHub |
| بسیار مناسب | معیار ارزیابی rank 1 : 75 % | ورودی این شبکه تصاویر رنگی به طول 224 در 224 می­باشد. | [link](https://github.com/elb3k/vtn) | VTN | [UCF-101](https://www.crcv.ucf.edu/data/UCF101.php) | 7 |
| معیار ارزیابی rank 1 : 79 % | [link](https://github.com/bomri/SlowFast/blob/master/projects/vtn/README.md) |
| نامناسب – کد ها تماما به زبان matlab است | 87.6% | - | ResNeXt-50 | [link](https://github.com/hbilen/dynamic-image-nets) | Dynamic Image Networks for Action Recognition | 8 |
| [link](https://www.robots.ox.ac.uk/~vgg/publications/2016/Bilen16a/) | Article Page |
| - | - | [Spatial Net](https://graphdeeplearning.github.io/project/spatial-convnets/) | [link](https://github.com/wanglimin/TDD) | GitHub | Trajectory-Pooled Deep-Convolutional Descriptors | 9 |
| مناسب | 82.1% | - | Spatial ResNet | [link](https://github.com/jeffreyyihuang/two-stream-action-recognition) | two-stream-action-recognition | 10 |
| نامناسب – نامرتبط با پروژه | [link](http://relation.csail.mit.edu/demo_v2.mp4) | Temporal relational reasoning, the ability to link meaningful transformations of objects or entities over time, is a fundamental property of intelligent species. We introduce an effective and interpretable network module, the Temporal Relation Network (TRN), designed to learn and reason about temporal dependencies between video frames at multiple time scales. Video Link: [link](http://relation.csail.mit.edu/demo_v2.mp4) | CNN | [link](http://relation.csail.mit.edu/) | Article Page | Temporal Relational Reasoning in Videos | [Something-Something](https://developer.qualcomm.com/software/ai-datasets/something-something) | 11 |
| [link](https://github.com/zhoubolei/TRN-pytorch) | GitHub | 12 |
| نامناسب - با توجه به مجموعه داده، این روش برای شناسایی اشیا در خصوص رانندگی بدون سرنشین کاربرد دارد | - | - | [RNN](https://raw.githubusercontent.com/xumingze0308/TRN.pytorch/master/demo/network.jpg) | [link](https://github.com/xumingze0308/TRN.pytorch) | Temporal Recurrent Networks for Online Action Detection | [HDD](https://usa.honda-ri.com/hdd) | 13 |
| نامناسب – نیاز به آموزش | 70 % | - | [LSTR](https://raw.githubusercontent.com/amazon-science/long-short-term-transformer/main/demo/network.png) | [link](https://github.com/amazon-science/long-short-term-transformer) | LSTR | [THUMOS14](https://www.crcv.ucf.edu/THUMOS14/) | 14 |
| مناسب | 49% | - | [CNN](http://yjxiong.me/others/ssn/overview.jpg) | [link](https://github.com/yjxiong/action-detection#code-and-data-preparation) | SSN | 15 |
| [link](http://yjxiong.me/others/ssn/) | Article Page |
| متوسط – نیاز به آموزش مدل | [link](https://github.com/jiyanggao/CBR-results) | - | CNN | [link](https://github.com/jiyanggao/CBR) | GitHub | CBR | 16 |
| بسیار مناسب | [link](https://github.com/felixchenfy/Data-Storage/raw/master/EECS-433-Pattern-Recognition/recog_actions.gif) | - | Mobile-Net | [link](https://github.com/felixchenfy/Realtime-Action-Recognition) | Multi-person Real-time Action Recognition Based-on Human Skeleton | Custom Dataset Containing 1000 Videos – 6 Labels – 20 min | 17 |
| [link](https://github.com/felixchenfy/Data-Storage/raw/master/EECS-433-Pattern-Recognition/recog_actions2.gif) |
| مناسب | [link](https://github.com/TianzhongSong/Real-Time-Action-Recognition/blob/master/files/action2.gif) | - | Calculate on Points | [link](https://github.com/TianzhongSong/Real-Time-Action-Recognition) | Real-Time-Action-Recognition | - | 18 |