1. Attention Pooling for Action Recognition
2. Contextual Action Recognition with R\*CNN
3. Action Recognition in Still Images Using Word Embeddings from Natural Language Descriptions
4. Learning Discriminative Action And Context Representations For Action Recognition In Still Images
5. Action Recognition in Still Images with Minimum Annotation Efforts
6. Scale coding bag of deep features for human attribute and action recognition
7. A Zero-Shot Architecture for Action Recognition in Still Images
8. Learning to Detect Human-Object Interactions
9. ICAN: Instance-Centric Attention Network for Human-Object Interaction Detection
10. Pairwise Body-Part Attention for Recognizing Human-Object Interactions
11. An Attention Model for group-level emotion recognition
12. Residual Attention Network for Image Classification
13. Densely Connected CNN with Multi-scale Feature Attention for Text Classification
14. Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering
15. R-FCN: Object Detection via Region-based Fully Convolutional Networks
16. Deep Visual Attention Prediction
17. Weakly-supervised learning of visual relations
18. Expanded Parts Model for Semantic Description of Humans in Still Images
19. Deep Visual Words: Improved Fisher Vector for Image Classification
20. Learning Conditioned Graph Structures for Interpretable Visual Question Answering
21. Single Shot Multibox Detector
22. Receptive Field Block Net for Accurate and Fast Object Detection
23. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention
24. CBAM: Convolutional Block Attention Module
25. Paying More Attention to Attention: Improving The Performance of Convolutional Neural
26. Networks via Attention Transfer
27. Progressive Attention Guided Recurrent Network for Salient Object Detection
28. DetNet: A Backbone network for Object Detection