www.zhuanzhi.ai

计算机视觉（computer vision） 荟萃

初步版本，水平有限，有错误或者不完善的地方，欢迎大家提建议和补充，会一直保持更新

- 进阶文章
  - **Video Classification**
  - **Object Tracking**
  - **Segmentation**
  - **Object Recognition**
  - **Image Captioning**
  - **Video Captioning**
  - **Visual Question Answering**
  - **Edge Detection**
  - **Human Pose Estimation**
  - **Image Generation**
- 课程
- 综述
- **Turorial**
- 图书
- 相关期刊与会议
  - 国际会议
  - 期刊
- 领域专家
  - 华人机构和学者
  - **North America**
  - **Europe**
  - **Australia**
  - **Asia and Middle East**
- **Software**
- **Datasets**
  - **Detection**
  - **Classification**
  - **Recognition**
  - **Tracking**
  - **Segmentation**
  - **Foreground/Background**
  - **Saliency Detection (source)**
  - **Video Surveillance**
  - **Multiview**

    - **Action**
    - **Human pose/Expression**
    - **Medical**
    - **Misc**
- **Challenge**
- **创业公司**
- **公众号**

入门学习

1. 计算机视觉：让冰冷的机器看懂这个多彩的世界 by 孙剑
    - http://www.msra.cn/zh-cn/news/features/computer-vision-20150210
2. UCLA 朱松纯: 正本清源·初探计算机视觉的三个源头、兼谈人工智能
    - https://mp.weixin.qq.com/s/2ytV5Bt50yhYOFYXYQe6ZQ
3. 深度学习与视觉计算 by 王亮 中科院自动化所
    - http://www.caai.cn/index.php?s=/Home/Article/qikande-tail/year/2017/month/04.html
4. 如何做好计算机视觉的研究？ by 微软 华刚博士
    - http://www.msra.cn/zh-cn/news/features/do-research-in-computer-vision-20161205
5. 计算机视觉 微软亚洲研究院系列文章
    - 通俗介绍计算机视觉在生活中的各种应用。
    - http://www.msra.cn/zh-cn/research/computer-vision
6. 计算机视觉随谈
    - http://blog.csdn.net/zouxy09/article/details/38639349
7. 计算机视觉：就在你我身边 微软

   https://mp.weixin.qq.com/s/rgvQeW9CwswbmcAI4BISNQ
8. 什么是计算机视觉？什么是机器视觉？

   https://mp.weixin.qq.com/s/PVom2BwEUXw3z68cra9xNQ
9. 卷积神经网络如何进行图像识别

- http://www.infoq.com/cn/articles/convolutional-neural-networks-image-recognition

10. 相似图片搜索的原理 阮一峰
    - http://www.ruanyifeng.com/blog/2011/07/principle_of_similar_image_search.html
11. 如何识别图像边缘？ 阮一峰
    - http://www.ruanyifeng.com/blog/2016/07/edge-recognition.html
12. 图像目标检测（Object Detection）原理与实现 （1-6）
    - http://www.voidcn.com/article/p-xnjyqlkj-ua.html
13. 运动目标跟踪系列（1-17）
    - http://blog.csdn.net/App*12062011/article/category/6269524/1*
14. 看图说话的 AI 小朋友——图像标注趣谈(上，下)
    - https://zhuanlan.zhihu.com/p/22408033
    - https://zhuanlan.zhihu.com/p/22520434
15. Video Analysis 相关领域介绍之 Video Captioning(视频 to 文字描述)
    - https://zhuanlan.zhihu.com/p/26730181
16. 从特斯拉到计算机视觉之「图像语义分割」
    - https://zhuanlan.zhihu.com/p/21824299
17. 计算机视觉识别简史：从 AlexNet、ResNet 到 Mask RCNN
    - https://github.com/Nikasa1889/HistoryObjectRecognition
    - https://mp.weixin.qq.com/s/ZKMi4gRfDRcTxzKlTQb-Mw
    - https://github.com/Nikasa1889/HistoryObjectRecognition/blob/master/HistoryOfObjectRecognition%20-%20A0.pdf
18. 深度学习在计算机视觉领域的前沿进展
    - https://zhuanlan.zhihu.com/p/24699780
19. 深度学习时代的计算机视觉

    https://zhuanlan.zhihu.com/p/24699780
20. 视觉求索 公众号相关文章系列，

- 浅谈人工智能：现状、任务、构架与统一 | 正本清源
- http://mp.weixin.qq.com/s/-wSYLu-XvOrsST8*KEUa-Q*
- 人生若只如初见 | 学术人生
- https://mp.weixin.qq.com/s/kFA7bI*FFjZQkBNDvcn01g*
- 初探计算机视觉的三个源头、兼谈人工智能丨正本清源
- https://mp.weixin.qq.com/s/2ytV5Bt50yhYOFYXYQe6ZQ

21. 深度学习大讲堂 公众号相关文章系列
    - 深度学习在目标跟踪中的应用 https://zhuanlan.zhihu.com/p/22334661
    - 深度学习在图像取证中的进展与趋势 https://zhuanlan.zhihu.com/p/23341157
    - 行人检测、跟踪与检索领域年度进展报告 https://zhuanlan.zhihu.com/p/26807041
    - 基于深度学习的目标检测研究进展 https://zhuanlan.zhihu.com/p/21412911
    - 基于深度学习的视觉实例搜索研究进展 https://zhuanlan.zhihu.com/p/22265265
    - 基于深度学习的 VQA（视觉问答）技术 https://zhuanlan.zhihu.com/p/22530291
    - 人脸识别简史与近期进展 https://zhuanlan.zhihu.com/p/21465605
    - 边缘检测领域年度进展报告 https://zhuanlan.zhihu.com/p/26848831
    - 目标跟踪领域进展报告 https://zhuanlan.zhihu.com/p/27293523

进阶论文

Image Classification

1. Microsoft Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition
   http://arxiv.org/pdf/1512.03385v1.pdf
   [http://image-net.org/challenges/talks/ilsvrc2015*deep*residual*learning*kaiminghe.pdf]

2. Microsoft Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, http://arxiv.org/pdf/1502.01852
3. Batch Normalization Sergey Ioffe, Christian Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shifthttp://arxiv.org/pdf/1502.03167
4. GoogLeNet Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, CVPR, 2015.
   http://arxiv.org/pdf/1409.4842
5. VGG-Net Karen Simonyan and Andrew Zisserman, Very Deep Convolutional Networks for Large-Scale Visual Recognition, ICLR, 2015.
   http://www.robots.ox.ac.uk/
   http://arxiv.org/pdf/1409.1556
6. AlexNet Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, NIPS, 2012. http://papers.nips.cc/book/advances-in-neural-information-processing-systems-25-2012

Object Detection

进阶文章

1. Deep Neural Networks for Object Detection （基于 DNN 的对象检测）NIPS2013:
   - https://cis.temple.edu/
2. R-CNN Rich feature hierarchies for accurate object detection and semantic segmentation:
   - https://arxiv.org/abs/1311.2524
3. Fast R-CNN :

- http://arxiv.org/abs/1504.08083

4. Faster R-CNN Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks:
   - http://arxiv.org/abs/1506.01497
5. Scalable Object Detection using Deep Neural Networks
   - http://arxiv.org/abs/1312.2249
6. Scalable, High-Quality Object Detection
   - http://arxiv.org/abs/1412.1441
7. SPP-Net Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition
   - http://arxiv.org/abs/1406.4729
8. DeepID-Net DeepID-Net: Deformable Deep Convolutional Neural Networks for Object Detection
   - http://www.ee.cuhk.edu.hk/%CB%9Cwlouyang/projects/imagenetDeepId/index.html
9. Object Detectors Emerge in Deep Scene CNNs
   - http://arxiv.org/abs/1412.6856
10. segDeepM: Exploiting Segmentation and Context in Deep Neural Networks for Object Detection
    - https://arxiv.org/abs/1502.04275
11. Object Detection Networks on Convolutional Feature Maps
    - http://arxiv.org/abs/1504.06066
12. Improving Object Detection with Deep Convolutional Networks via Bayesian Optimization and Structured Prediction
    - http://arxiv.org/abs/1504.03293
13. DeepBox: Learning Objectness with Convolutional Networks
    - http://arxiv.org/abs/1504.03293
14. Object detection via a multi-region & semantic segmentation-aware CNN model

- http://arxiv.org/abs/1505.01749

15. You Only Look Once: Unified, Real-Time Object Detection
    - http://arxiv.org/abs/1506.02640
16. YOLOv2 YOLO9000: Better, Faster, Stronger
    - https://arxiv.org/abs/1612.08242
17. AttentionNet: Aggregating Weak Directions for Accurate Object Detection
    - http://arxiv.org/abs/1506.07704
18. DenseBox: Unifying Landmark Localization with End to End Object Detection
    - http://arxiv.org/abs/1509.04874
19. SSD: Single Shot MultiBox Detector
    - http://arxiv.org/abs/1512.02325
20. DSSD : Deconvolutional Single Shot Detector
    - https://arxiv.org/abs/1701.06659
21. G-CNN: an Iterative Grid Based Object Detector
    - http://arxiv.org/abs/1512.07729
22. HyperNet: Towards Accurate Region Proposal Generation and Joint Object Detection
    - http://arxiv.org/abs/1604.00600
23. A MultiPath Network for Object Detection
    - http://arxiv.org/abs/1604.02135
24. R-FCN: Object Detection via Region-based Fully Convolutional Networks
    - http://arxiv.org/abs/1605.06409
25. A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection
    - http://arxiv.org/abs/1607.07155
26. PVANET: Deep but Lightweight Neural Networks for Real-time Object Detection
    - http://arxiv.org/abs/1608.08021
27. Feature Pyramid Networks for Object Detection
    - https://arxiv.org/abs/1612.03144
28. Learning Chained Deep Features and Classifiers for Cascade in Object Detection

- https://arxiv.org/abs/1702.07054

29. DSOD: Learning Deeply Supervised Object Detectors from Scratch
    - https://arxiv.org/abs/1708.01241
30. Focal Loss for Dense Object Detection ICCV 2017 Best student paper award. Facebook AI Research
    - https://arxiv.org/abs/1708.02002
31. Mask-RCNN 2017 ICCV 2017 Best paper award. Facebook AI Research
    - https://arxiv.org/pdf/1703.06870.pdf

Video Classification

1. Nicolas Ballas, Li Yao, Pal Chris, Aaron Courville, "Delving Deeper into Convolutional Networks for Learning Video Representations", ICLR 2016.
   http://arxiv.org/pdf/1511.06432v4.pdf
2. Michael Mathieu, camille couprie, Yann Lecun, "Deep Multi Scale Video Prediction Beyond Mean Square Error", ICLR 2016. Paper
   http://arxiv.org/pdf/1511.05440v6.pdf
3. Donahue, Jeffrey, et al. Long-term recurrent convolutional networks for visual recognition and description CVPR 2015
   https://arxiv.org/abs/1411.4389
4. Karpathy, Andrej, et al. Large-scale Video Classification with Convolutional Neural Networks. CVPR 2014
   http://cs.stanford.edu/people/karpathy/deepvideo/
5. Yue-Hei Ng, Joe, et al. Beyond short snippets: Deep networks for video classification. CVPR 2015
   https://arxiv.org/abs/1503.08909
6. Tran, Du, et al. Learning Spatiotemporal Features with 3D Convolutional Networks. ICCV 2015
   https://arxiv.org/abs/1412.0767

Object Tracking

NIPS2013

- DLT: Naiyan Wang and Dit-Yan Yeung.

"Learning A Deep Compact Image Representation for Visual Tracking." NIPS (2013).

- paper http://winsty.net/papers/dlt.pdf)
- project http://winsty.net/dlt.html)
- code http://winsty.net/dlt/DLTcode.zip)

CVPR2014

- CN: Martin Danelljan, Fahad Shahbaz Khan, Michael Felsberg and Joost van de Weijer.

"Adaptive Color Attributes for Real-Time Visual Tracking." CVPR (2014).

- paper http://www.cvl.isy.liu.se/research/objrec/visualtracking/colvistrack/CN*Track-ing*CVPR14.pdf
- project http://www.cvl.isy.liu.se/research/objrec/visualtracking/colvistrack/index.html

ECCV2014

- MEEM: Jianming Zhang, Shugao Ma, and Stan Sclaroff.

"MEEM: Robust Tracking via Multiple Experts using Entropy Minimization." ECCV (2014).

- paper http://cs-people.bu.edu/jmzhang/MEEM/MEEM-eccv-preprint.pdf
- project http://cs-people.bu.edu/jmzhang/MEEM/MEEM.html
- TGPR: Jin Gao, Haibin Ling, Weiming Hu, Junliang Xing.

"Transfer Learning Based Visual Tracking with Gaussian Process Regression." ECCV (2014).

- paper http://www.dabi.temple.edu/
- project http://www.dabi.temple.edu/

- STC: Kaihua Zhang, Lei Zhang, Ming-Hsuan Yang, David Zhang.

"Fast Tracking via Spatio-Temporal Context Learning." ECCV (2014).

- paper http://arxiv.org/pdf/1311.1939v1.pdf
- project http://www4.comp.polyu.edu.hk/
- SAMF: Yang Li, Jianke Zhu.

"A Scale Adaptive Kernel Correlation Filter Tracker with Feature Integration." ECCV workshop (2014).

- paper http://link.springer.com/content/pdf/10.1007%2F978-3-319-16181-5*18.pdf*
- github https://github.com/ihpdep/samf

BMVC2014

- DSST: Martin Danelljan, Gustav Häger, Fahad Shahbaz Khan and Michael Felsberg.

"Accurate Scale Estimation for Robust Visual Tracking." BMVC (2014).

- paper http://www.cvl.isy.liu.se/research/objrec/visualtracking/scalvistrack/ScaleTracking*BMVC14.pdf*
- PAMI http://www.cvl.isy.liu.se/en/research/objrec/visualtracking/scalvistrack/DSST*TPAMI.pdf*
- project http://www.cvl.isy.liu.se/en/research/objrec/visualtracking/scalvistrack/index.html
- SAMF: Yang Li, Jianke Zhu.

"A Scale Adaptive Kernel Correlation Filter Tracker with Feature Integration." ECCV workshop (2014).

- paper https://github.com/ihpdep/ihpdep.github.io/raw/master/papers/eccvw14*samf.pdf*
- github https://github.com/ihpdep/samf

ICML2015

- CNN-SVM: Seunghoon Hong, Tackgeun You, Suha Kwak and Bohyung Han.

"Online Tracking by Learning Discriminative Saliency Map with Convolutional Neural Network ." ICML (2015)

- paper http://120.52.73.80/arxiv.org/pdf/1502.06796.pdf
- project http://cvlab.postech.ac.kr/research/CNN*SVM/*

CVPR2015

- MUSTer: Zhibin Hong, Zhe Chen, Chaohui Wang, Xue Mei, Danil Prokhorov, Dacheng Tao.

"MUlti-Store Tracker (MUSTer): A Cognitive Psychology Inspired Approach to Object Tracking." CVPR (2015).

- paper http://openaccess.thecvf.com/content*cvpr*2015/papers/Hong*MUlti-Store*Tracker-*MUSTer*2015*CVPR*paper.pdf
- project https://sites.google.com/site/multistoretrackermuster/
- LCT: Chao Ma, Xiaokang Yang, Chongyang Zhang, Ming-Hsuan Yang.

"Long-term Correlation Tracking." CVPR (2015).

- paper http://openaccess.thecvf.com/content*cvpr*2015/papers/Ma*Long-Term*Correlation*Tracking*2015*CVPR*paper.pdf
- project https://sites.google.com/site/chaoma99/cvpr15*tracking*
- github https://github.com/chaoma99/lct-tracker
- DAT: Horst Possegger, Thomas Mauthner, and Horst Bischof.

"In Defense of Color-based Model-free Tracking." CVPR (2015).

- paper https://lrs.icg.tugraz.at/pubs/possegger*cvpr15.pdf*

- project https://www.tugraz.at/institute/icg/research/team-bischof/lrs/downloads/dat
- code https://lrs.icg.tugraz.at/downloads/dat-v1.0.zip
- RPT: Yang Li, Jianke Zhu and Steven C.H. Hoi.

"Reliable Patch Trackers: Robust Visual Tracking by Exploiting Reliable Patches." CVPR (2015).

- paper https://github.com/ihpdep/ihpdep.github.io/raw/master/papers/cvpr15*rpt.pdf*
- github https://github.com/ihpdep/rpt

ICCV2015

- FCNT: Lijun Wang, Wanli Ouyang, Xiaogang Wang, and Huchuan Lu.

"Visual Tracking with Fully Convolutional Networks." ICCV (2015).

- paper http://202.118.75.4/lu/Paper/ICCV2015/iccv15*lijun.pdf*
- project http://scott89.github.io/FCNT/
- github https://github.com/scott89/FCNT
- SRDCF: Martin Danelljan, Gustav Häger, Fahad Khan, Michael Felsberg.

"Learning Spatially Regularized Correlation Filters for Visual Tracking." ICCV (2015).

- paper https://www.cvl.isy.liu.se/research/objrec/visualtracking/regvistrack/SRDCF*ICCV15.pdf*
- project https://www.cvl.isy.liu.se/research/objrec/visualtracking/regvistrack/
- CF2: Chao Ma, Jia-Bin Huang, Xiaokang Yang and Ming-Hsuan Yang.

"Hierarchical Convolutional Features for Visual Tracking." ICCV (2015)

- paper http://faculty.ucmerced.edu/mhyang/papers/iccv15*tracking.pdf*
- project https://sites.google.com/site/jbhuang0604/publications/cf2
- github https://github.com/jbhuang0604/CF2

- Naiyan Wang, Jianping Shi, Dit-Yan Yeung and Jiaya Jia.

"Understanding and Diagnosing Visual Tracking Systems." ICCV (2015).

- paper http://winsty.net/papers/diagnose.pdf
- project http://winsty.net/tracker*diagnose.html*
- code http://winsty.net/diagnose/diagnose*code.zip*
- DeepSRDCF: Martin Danelljan, Gustav Häger, Fahad Khan, Michael Felsberg.

"Convolutional Features for Correlation Filter Based Visual Tracking." ICCV workshop (2015).

- paper https://www.cvl.isy.liu.se/research/objrec/visualtracking/regvistrack/ConvDCF*ICCV15*VOTworkshop.pdf
- project https://www.cvl.isy.liu.se/research/objrec/visualtracking/regvistrack/
- RAJSSC: Mengdan Zhang, Junliang Xing, Jin Gao, Xinchu Shi, Qiang Wang, Weiming Hu.

"Joint Scale-Spatial Correlation Tracking with Adaptive Rotation Estimation." ICCV workshop (2015).

- paper http://www.cv-foundation.org//openaccess/content*iccv*2015*workshops/w14/papers/Zhang*Joint*Scale-Spatial*Correlation*ICCV*2015*paper.pdf*
- poster http://www.votchallenge.net/vot2015/download/poster*Mengdan*Zhang.pdf

NIPS2016

- Learnet: Luca Bertinetto, João F. Henriques, Jack Valmadre, Philip H. S. Torr, Andrea Vedaldi.

"Learning feed-forward one-shot learners." NIPS (2016).

- paper https://arxiv.org/pdf/1606.05233v1.pdf

CVPR2016

- MDNet: Nam, Hyeonseob, and Bohyung Han.

"Learning Multi-Domain Convolutional Neural Networks for Visual Tracking." CVPR (2016).

- paper http://arxiv.org/pdf/1510.07945v2.pdf
- VOT*presentation* *http://votchallenge.net/vot2015/download/presentation*Hyeonseob.pdf]
- project http://cvlab.postech.ac.kr/research/mdnet/
- github https://github.com/HyeonseobNam/MDNet
- SINT: Ran Tao, Efstratios Gavves, Arnold W.M. Smeulders.

"Siamese Instance Search for Tracking." CVPR (2016).

- paper https://staff.science.uva.nl/r.tao/pub/TaoCVPR2016.pdf
- project https://staff.fnwi.uva.nl/r.tao/projects/SINT/SINT*proj.html*
- SCT: Jongwon Choi, Hyung Jin Chang, Jiyeoup Jeong, Yiannis Demiris, and Jin Young Choi.

"Visual Tracking Using Attention-Modulated Disintegration and Integration." CVPR (2016).

- paper http://www.cv-foundation.org/openaccess/content*cvpr*2016/papers/Choi*Visual*-Tracking*Using*CVPR*2016*paper.pdf
- project https://sites.google.com/site/jwchoivision/home/sct
- STCT: Lijun Wang, Wanli Ouyang, Xiaogang Wang, and Huchuan Lu.

"STCT: Sequentially Training Convolutional Networks for Visual Tracking." CVPR (2016).

- paper http://www.ee.cuhk.edu.hk/
- github https://github.com/scott89/STCT
- SRDCFdecon: Martin Danelljan, Gustav Häger, Fahad Khan, Michael Felsberg.

"Adaptive Decontamination of the Training Set: A Unified Formulation for Discriminative Visual Tracking." CVPR (2016).

- paper https://www.cvl.isy.liu.se/research/objrec/visualtracking/decontrack/AdaptiveDecon*CVPR16.pdf*
- project https://www.cvl.isy.liu.se/research/objrec/visualtracking/decontrack/index.html
- HDT: Yuankai Qi, Shengping Zhang, Lei Qin, Hongxun Yao, Qingming Huang, Jongwoo Lim, Ming-Hsuan Yang.

"Hedged Deep Tracking." CVPR (2016).

- paper http://faculty.ucmerced.edu/mhyang/papers/cvpr16*hedge*tracking.pdf
- project https://sites.google.com/site/yuankiqi/hdt/
- Staple: Luca Bertinetto, Jack Valmadre, Stuart Golodetz, Ondrej Miksik, Philip H.S. Torr.

"Staple: Complementary Learners for Real-Time Tracking." CVPR (2016).

- paper http://120.52.73.75/arxiv.org/pdf/1512.01355v2.pdf
- project http://www.robots.ox.ac.uk/
- github https://github.com/bertinetto/staple
- EBT: Gao Zhu, Fatih Porikli, and Hongdong Li.

"Beyond Local Search: Tracking Objects Everywhere with Instance-Specific Proposals." CVPR (2016).

- paper http://www.cv-foundation.org/openaccess/content*cvpr*2016/papers/Zhu-*Beyond*Local*Search*CVPR*2016*paper.pdf
- exe http://www.votchallenge.net/vot2016/download/02*EBT.zip*
- DLSSVM: Jifeng Ning, Jimei Yang, Shaojie Jiang, Lei Zhang and Ming-Hsuan Yang.

"Object Tracking via Dual Linear Structured SVM and Explicit Feature Map." CVPR (2016).

- paper http://www4.comp.polyu.edu.hk/
- code http://www4.comp.polyu.edu.hk/
- project http://www4.comp.polyu.edu.hk/

ECCV2016

- SiameseFC: Luca Bertinetto, Jack Valmadre, João F. Henriques, Andrea Vedaldi, Philip H.S. Torr.

"Fully-Convolutional Siamese Networks for Object Tracking." ECCV workshop (2016).

- paper http://120.52.73.78/arxiv.org/pdf/1606.09549v2.pdf
- project http://www.robots.ox.ac.uk/
- github https://github.com/bertinetto/siamese-fc
- GOTURN: David Held, Sebastian Thrun, Silvio Savarese.

"Learning to Track at 100 FPS with Deep Regression Networks." ECCV (2016).

- paper http://davheld.github.io/GOTURN/GOTURN.pdf
- project http://davheld.github.io/GOTURN/GOTURN.html
- github https://github.com/davheld/GOTURN
- C-COT: Martin Danelljan, Andreas Robinson, Fahad Khan, Michael Felsberg.

"Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking." ECCV (2016).

- paper http://www.cvl.isy.liu.se/research/objrec/visualtracking/conttrack/C-COT*ECCV16.pdf*
- project http://www.cvl.isy.liu.se/research/objrec/visualtracking/conttrack/index.html
- github https://github.com/martin-danelljan/Continuous-ConvOp
- CF+AT: Adel Bibi, Matthias Mueller, and Bernard Ghanem.

"Target Response Adaptation for Correlation Filter Tracking." ECCV (2016).

- paper http://www.adelbibi.com/papers/ECCV2016/Target*Adap.pdf*
- Yao Sui, Ziming Zhang, Guanghui Wang, Yafei Tang, Li Zhang.

"Real-Time Visual Tracking: Promoting the Robustness of Correlation Filter Learning." ECCV (2016).

- paper http://120.52.73.78/arxiv.org/pdf/1608.08173.pdf
- Yao Sui, Guanghui Wang, Yafei Tang, Li Zhang.

"Tracking Completion." ECCV (2016).

- paper http://120.52.73.78/arxiv.org/pdf/1608.08171v1.pdf

CVPR2017

- ECO: Martin Danelljan, Goutam Bhat, Fahad Shahbaz Khan, Michael Felsberg.

"ECO: Efficient Convolution Operators for Tracking." CVPR (2017).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Danelljan*ECO*Efficient*Convolution*CVPR*2017*paper.pdf
- supp http://openaccess.thecvf.com/content*cvpr*2017/supplemental/Danelljan*ECO*Efficient*Convolution*2017*CVPR*supplemental.pdf
- project http://www.cvl.isy.liu.se/research/objrec/visualtracking/ecotrack/index.html
- github https://github.com/martin-danelljan/ECO
- CFNet: Jack Valmadre, Luca Bertinetto, João F. Henriques, Andrea Vedaldi, Philip H. S. Torr.

"End-to-end representation learning for Correlation Filter based tracking." CVPR (2017).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Valmadre*End-To-End*Representation*Learning*CVPR*2017*paper.pdf

- supp http://openaccess.thecvf.com/content*cvpr*2017/supplemental/Valmadre*End-To-End*Representation*Learning*2017*CVPR*supplemental.pdf
- project http://www.robots.ox.ac.uk/
- github https://github.com/bertinetto/cfnet
- CACF: Matthias Mueller, Neil Smith, Bernard Ghanem.

"Context-Aware Correlation Filter Tracking." CVPR (2017 oral).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Mueller*Context-Aware*Correlation*Filter*CVPR*2017*paper.pdf
- supp http://openaccess.thecvf.com/content*cvpr*2017/supplemental/Mueller*Context-Aware*Correlation*Filter*2017*CVPR*supplemental.zip
- project https://ivul.kaust.edu.sa/Pages/pub-ca-cf-tracking.aspx
- code https://github.com/thias15/Context-Aware-CF-Tracking
- RaF: Le Zhang, Jagannadan Varadarajan, Ponnuthurai Nagaratnam Suganthan, Narendra Ahuja and Pierre Moulin

"Robust Visual Tracking Using Oblique Random Forests." CVPR (2017).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Zhang*Robust*Visual*Tracking*CVPR*2017*paper.pdf
- supp http://openaccess.thecvf.com/content*cvpr*2017/supplemental/Zhang*Robust*Visual*Tracking*2017*CVPR*supplemental.pdf
- project https://sites.google.com/site/zhangleuestc/incremental-oblique-random-forest
- code https://github.com/ZhangLeUestc/Incremental-Oblique-Random-Forest
- MCPF: Tianzhu Zhang, Changsheng Xu, Ming-Hsuan Yang.

"Multi-Task Correlation Particle Filter for Robust Object Tracking." CVPR (2017).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Zhang*Multi-Task*Correlation*Particle*CVPR*2017*paper.pdf

- project http://nlpr-web.ia.ac.cn/mmc/homepage/tzzhang/mcpf.html
- code http://nlpr-web.ia.ac.cn/mmc/homepage/tzzhang/mcpf.html
- ACFN: Jongwon Choi, Hyung Jin Chang, Sangdoo Yun, Tobias Fischer, Yiannis Demiris, and Jin Young Choi.

"Attentional Correlation Filter Network for Adaptive Visual Tracking." CVPR (2017).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Choi*Attentional*Correlation*Filter*CVPR*2017*paper.pdf
- supp http://openaccess.thecvf.com/content*cvpr*2017/supplemental/Choi*Attentional*Correlation*Filter*2017*CVPR*supplemental.pdf
- project https://sites.google.com/site/jwchoivision/home/acfn-1
- test code https://drive.google.com/file/d/0B0ZkG8zaRQoLQUswbW9qSWFaU0U/view?usp=drive*web*
- training code https://drive.google.com/file/d/0B0ZkG8zaRQoLZVVranBnbHlydnM/view?usp=drive*web*
- LMCF: Mengmeng Wang, Yong Liu, Zeyi Huang.

"Large Margin Object Tracking with Circulant Feature Maps." CVPR (2017).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Wang*Large*Margin*Object*CVPR*2017*paper.pdf
- zhihu https://zhuanlan.zhihu.com/p/25761718
- ADNet: Sangdoo Yun, Jongwon Choi, Youngjoon Yoo, Kimin Yun, Jin Young Choi.

"Action-Decision Networks for Visual Tracking with Deep Reinforcement Learning." CVPR (2017 Spotlight).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Yun*Action-Decision*Networks*for*CVPR*2017*paper.pdf

- supp http://openaccess.thecvf.com/content*cvpr*2017/supplemental/Yun*Action-Decision*Networks*for*2017*CVPR*supplemental.pdf
- project https://sites.google.com/view/cvpr2017-adnet
- CSR-DCF: Alan Lukežič, Tomáš Vojíř, Luka Čehovin, Jiří Matas, Matej Kristan.

"Discriminative Correlation Filter with Channel and Spatial Reliability." CVPR (2017).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Lukezic*Discriminative*Correlation*Filter*CVPR*2017*paper.pdf
- supp http://openaccess.thecvf.com/content*cvpr*2017/supplemental/Lukezic*Discriminative*Correlation*Filter*2017*CVPR*supplemental.pdf
- code https://github.com/alanlukezic/csr-dcf
- BranchOut: Bohyung Han, Jack Sim, Hartwig Adam.

"BranchOut: Regularization for Online Ensemble Tracking with Convolutional Neural Networks." CVPR (2017).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Han*BranchOut*Regularization*for*CVPR*2017*paper.pdf
- AMCT: Donghun Yeo, Jeany Son, Bohyung Han, Joonhee Han.

"Superpixel-based Tracking-by-Segmentation using Markov Chains." CVPR (2017).

- paper http://openaccess.thecvf.com/content*cvpr*2017/papers/Yeo*Superpixel-Based-Tracking-By-Segmentation*Using*CVPR*2017*paper.pdf
- SANet: Heng Fan, Haibin Ling.

"SANet: Structure-Aware Network for Visual Tracking." CVPRW (2017).

- paper https://arxiv.org/pdf/1611.06878.pdf
- project http://www.dabi.temple.edu/
- code http://www.dabi.temple.edu/

ICCV2017

- CREST: Yibing Song, Chao Ma, Lijun Gong, Jiawei Zhang, Rynson Lau, Ming-Hsuan Yang.

"CREST: Convolutional Residual Learning for Visual Tracking." ICCV (2017 Spotlight).

- paper http://openaccess.thecvf.com/content*ICCV*2017/papers/Song*CREST*Convolutional*Residual*ICCV*2017*paper.pdf
- project http://www.cs.cityu.edu.hk/
- github https://github.com/ybsong00/CREST-Release
- EAST: Chen Huang, Simon Lucey, Deva Ramanan.

"Learning Policies for Adaptive Tracking with Deep Feature Cascades." ICCV (2017 Spotlight).

- paper http://openaccess.thecvf.com/content*ICCV*2017/papers/Huang*Learning*Policies*for*ICCV*2017*paper.pdf
- supp http://openaccess.thecvf.com/content*ICCV*2017/supplemental/Huang*Learning*Policies*for*ICCV*2017*supplemental.zip
- PTAV: Heng Fan and Haibin Ling.

"Parallel Tracking and Verifying: A Framework for Real-Time and High Accuracy Visual Tracking." ICCV (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017/papers/Fan*Parallel*Tracking*and*ICCV*2017*paper.pdf
- supp http://openaccess.thecvf.com/content*ICCV*2017/supplemental/Fan*Parallel*Tracking*and*ICCV*2017*supplemental.pdf
- project http://www.dabi.temple.edu/
- code http://www.dabi.temple.edu/
- BACF: Hamed Kiani Galoogahi, Ashton Fagg, Simon Lucey.

"Learning Background-Aware Correlation Filters for Visual Tracking." ICCV (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017/papers/Galoogahi*Learning*-Background-Aware*Correlation*ICCV*2017*paper.pdf
- supp http://openaccess.thecvf.com/content*ICCV*2017/supplemental/Galoogahi*Learning*Background-Aware*Correlation*ICCV*2017*supplemental.pdf
- TSN: Zhu Teng, Junliang Xing, Qiang Wang, Congyan Lang, Songhe Feng and Yi Jin.

"Robust Object Tracking based on Temporal and Spatial Deep Networks." ICCV (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017/papers/Teng*Robust*Object*Tracking*ICCV*2017*paper.pdf
- p-tracker: James Supančič, III; Deva Ramanan.

"Tracking as Online Decision-Making: Learning a Policy From Streaming Videos With Reinforcement Learning." ICCV (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017/papers/Supancic*Tracking*as*Online*ICCV*2017*paper.pdf
- supp http://openaccess.thecvf.com/content*ICCV*2017/supplemental/Supancic*Tracking*as*Online*ICCV*2017*supplemental.pdf
- DSiam: Qing Guo; Wei Feng; Ce Zhou; Rui Huang; Liang Wan; Song Wang.

"Learning Dynamic Siamese Network for Visual Object Tracking." ICCV (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017/papers/Guo*Learning*Dynamic-*Siamese*ICCV*2017*paper.pdf
- SP-KCF: Xin Sun; Ngai-Man Cheung; Hongxun Yao; Yiluan Guo.

"Non-Rigid Object Tracking via Deformable Patches Using Shape-Preserved KCF and Level Sets." ICCV (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017/papers/Sun*Non-Rigid*Object-*Tracking*ICCV*2017*paper.pdf
- UCT: Zheng Zhu, Guan Huang, Wei Zou, Dalong Du, Chang Huang.

"UCT: Learning Unified Convolutional Networks for Real-Time Visual Tracking." ICCV workshop (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017*workshops/papers/w28/Zhu*UCT*Learning*Unified*ICCV*2017*paper.pdf*
- Tobias Bottger, Patrick Follmann.

"The Benefits of Evaluating Tracker Performance Using Pixel-Wise Segmentations." ICCV workshop (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017*workshops/papers/w28/Bottger-*The*Benefits*of*ICCV*2017*paper.pdf*
- CFWCR: Zhiqun He, Yingruo Fan, Junfei Zhuang, Yuan Dong, HongLiang Bai.

"Correlation Filters With Weighted Convolution Responses." ICCV workshop (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017*workshops/papers/w28/He*Correlation*Filters*With*ICCV*2017*paper.pdf*
- IBCCF: Feng Li, Yingjie Yao, Peihua Li, David Zhang, Wangmeng Zuo, Ming-Hsuan Yang.

"Integrating Boundary and Center Correlation Filters for Visual Tracking With Aspect Ratio Variation." ICCV workshop (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017*workshops/papers/w28/Li*Integrating*Boundary*and*ICCV*2017*paper.pdf*
- RFL: Tianyu Yang, Antoni B. Chan.

"Recurrent Filter Learning for Visual Tracking." ICCV workshop (2017).

- paper http://openaccess.thecvf.com/content*ICCV*2017*workshops/papers/w28/Yang*Re-current*Filter*Learning*ICCV*2017*paper.pdf*

PAMI & IJCV & TIP

- KCF: João F. Henriques, Rui Caseiro, Pedro Martins, Jorge Batista.

"High-Speed Tracking with Kernelized Correlation Filters." TPAMI (2015).

- paper http://www.robots.ox.ac.uk/
- project http://www.robots.ox.ac.uk/
- CLRST: Tianzhu Zhang, Si Liu, Narendra Ahuja, Ming-Hsuan Yang, Bernard Ghanem.

"Robust Visual Tracking Via Consistent Low-Rank Sparse Learning." IJCV (2015).

- paper http://nlpr-web.ia.ac.cn/mmc/homepage/tzzhang/tianzhu%20zhang*files/Jour-nal%20Articles/IJCV15*zhang*Low-Rank%20Sparse%20Learning.pdf*
- project http://nlpr-web.ia.ac.cn/mmc/homepage/tzzhang/Project*Tianzhu/zhang*I-JCV14/Robust%20Visual%20Tracking%20Via%20Consistent%20Low-Rank%20Sparse.html
- code http://nlpr-web.ia.ac.cn/mmc/homepage/tzzhang/Project*Tianzhu/zhang*I-JCV14/material/LRT*Code.zip*
- DNT: Zhizhen Chi, Hongyang Li, Huchuan Lu, Ming-Hsuan Yang.

"Dual Deep Network for Visual Tracking." TIP (2017).

- paper https://arxiv.org/pdf/1612.06053v1.pdf
- DRT: Junyu Gao, Tianzhu Zhang, Xiaoshan Yang, Changsheng Xu.

"Deep Relative Tracking." TIP (2017).

- paper http://ieeexplore.ieee.org/abstract/document/7828108/
- BIT: Bolun Cai, Xiangmin Xu, Xiaofen Xing, Kui Jia, Jie Miao, Dacheng Tao.

"BIT: Biologically Inspired Tracker." TIP (2016).

- paper http://caibolun.github.io/papers/BIT*TIP.pdf*
- project http://caibolun.github.io/BIT/index.html
- github https://github.com/caibolun/BIT
- CNT: Kaihua Zhang, Qingshan Liu, Yi Wu, Minghsuan Yang.

"Robust Visual Tracking via Convolutional Networks Without Training." TIP (2016).

- paper http://kaihuazhang.net/CNT.pdf
- code http://kaihuazhang.net/CNT*matlab.rar*

ArXiv

- PAWSS: Xiaofei Du, Alessio Dore, Danail Stoyanov.

"Patch-based adaptive weighting with segmentation and scale (PAWSS) for visual tracking." arXiv (2017).

- paper https://arxiv.org/pdf/1708.01179v1.pdf
- SFT: Zhen Cui, You yi Cai, Wen ming Zheng, Jian Yang.

"Spectral Filter Tracking." arXiv (2017).

- paper https://arxiv.org/pdf/1707.05553v1.pdf
- HART: Adam R. Kosiorek, Alex Bewley, Ingmar Posner.

"Hierarchical Attentive Recurrent Tracking." arXiv (2017).

- paper https://arxiv.org/pdf/1706.09262.pdf
- github https://github.com/akosiorek/hart
- Re3: Daniel Gordon, Ali Farhadi, Dieter Fox.

"Re3 : Real-Time Recurrent Regression Networks for Object Tracking." arXiv (2017).

- paper https://arxiv.org/pdf/1705.06368.pdf
- DCFNet: Qiang Wang, Jin Gao, Junliang Xing, Mengdan Zhang, Weiming Hu.

"DCFNet: Discriminant Correlation Filters Network for Visual Tracking." arXiv (2017).

- paper https://arxiv.org/pdf/1704.04057.pdf
- code https://github.com/foolwood/DCFNet#dcfnet-discriminant-correlation-filters-network-for-visual-tracking
- TCNN: Hyeonseob Nam, Mooyeol Baek, Bohyung Han.

"Modeling and Propagating CNNs in a Tree Structure for Visual Tracking." arXiv (2016).

- paper http://arxiv.org/pdf/1608.07242v1.pdf
- code http://www.votchallenge.net/vot2016/download/44*TCNN.zip*
- RDT: Janghoon Choi, Junseok Kwon, Kyoung Mu Lee.

"Visual Tracking by Reinforced Decision Making." arXiv (2017).

- paper https://arxiv.org/pdf/1702.06291.pdf
- MSDAT: Xinyu Wang, Hanxi Li, Yi Li, Fumin Shen, Fatih Porikli .

"Robust and Real-time Deep Tracking Via Multi-Scale Domain Adaptation." arXiv (2017).

- paper https://arxiv.org/pdf/1701.00561.pdf
- RLT: Da Zhang, Hamid Maei, Xin Wang, Yuan-Fang Wang.

"Deep Reinforcement Learning for Visual Object Tracking in Videos." arXiv (2017).

- paper](https://arxiv.org/pdf/1701.08936v1.pdf]
- SCF: Wangmeng Zuo, Xiaohe Wu, Liang Lin, Lei Zhang, Ming-Hsuan Yang.

"Learning Support Correlation Filters for Visual Tracking." arXiv (2016).

- paper https://arxiv.org/pdf/1601.06032.pdf

- project http://faculty.ucmerced.edu/mhyang/project/scf/
- CRT: Kai Chen, Wenbing Tao.

"Convolutional Regression for Visual Tracking." arXiv (2016).

- paper](https://arxiv.org/pdf/1611.04215.pdf]
- BMR: Kaihua Zhang, Qingshan Liu, and Ming-Hsuan Yang.

"Visual Tracking via Boolean Map Representations." arXiv (2016).

- paper](https://arxiv.org/pdf/1610.09652v1.pdf]
- YCNN: Kai Chen, Wenbing Tao.

"Once for All: a Two-flow Convolutional Neural Network for Visual Tracking." arXiv (2016).

- paper](https://arxiv.org/pdf/1604.07507v1.pdf]
- ROLO: Guanghan Ning, Zhi Zhang, Chen Huang, Zhihai He, Xiaobo Ren, Haohong Wang.

"Spatially Supervised Recurrent Convolutional Neural Networks for Visual Object Tracking." arXiv (2016).

- paper http://arxiv.org/pdf/1607.05781v1.pdf
- project http://guanghan.info/projects/ROLO/
- github https://github.com/Guanghan/ROLO/
- RATM: Samira Ebrahimi Kahou, Vincent Michalski, Roland Memisevic.

"RATM: Recurrent Attentive Tracking Model." arXiv (2015).

- paper https://arxiv.org/pdf/1510.08660v4.pdf
- github https://github.com/saebrahimi/RATM
- SO-DLT: Naiyan Wang, Siyi Li, Abhinav Gupta, Dit-Yan Yeung.

"Transferring Rich Feature Hierarchies for Robust Visual Tracking." arXiv (2015).

- paper https://arxiv.org/pdf/1501.04587v2.pdf
- code http://www.votchallenge.net/vot2016/download/08*SO-DLT.zip*
- DMSRDCF: Susanna Gladh, Martin Danelljan, Fahad Shahbaz Khan, Michael Felsberg.

"Deep Motion Features for Visual Tracking." ICPR Best Paper (2016).

- paper](https://arxiv.org/pdf/1612.06615v1.pdf]

Segmentation

1. Alexander Kolesnikov, Christoph Lampert, Seed, Expand and Constrain: Three Principles for Weakly-Supervised Image Segmentation, ECCV, 2016.
   http://pub.ist.ac.at/
   https://github.com/kolesman/SEC
2. Guosheng Lin, Chunhua Shen, Ian Reid, Anton van dan Hengel, Efficient piecewise training of deep structured models for semantic segmentation, arXiv:1504.01013.
   http://arxiv.org/pdf/1504.01013
3. Guosheng Lin, Chunhua Shen, Ian Reid, Anton van den Hengel, Deeply Learning the Messages in Message Passing Inference, arXiv:1508.02108.
   http://arxiv.org/pdf/1506.02108
4. Deep Parsing Network . Ziwei Liu, Xiaoxiao Li, Ping Luo, Chen Change Loy, Xiaoou Tang, Semantic Image Segmentation via Deep Parsing Network, arXiv:1509.02634 / ICCV 2015
   http://arxiv.org/pdf/1509.02634.pdf
5. CentraleSuperBoundaries, Iasonas Kokkinos, Surpassing Humans in Boundary Detection using Deep Learning INRIA http://arxiv.org/pdf/1511.07386
6. BoxSup. Jifeng Dai, Kaiming He, Jian Sun, BoxSup: Exploiting Bounding Boxes to Supervise Convolutional Networks for Semantic Segmentation
   http://arxiv.org/pdf/1503.01640

7. Hyeonwoo Noh, Seunghoon Hong, Bohyung Han, Learning Deconvolution Network for Semantic Segmentation, arXiv:1505.04366. http://arxiv.org/pdf/1505.04366
8. Seunghoon Hong, Hyeonwoo Noh, Bohyung Han, Decoupled Deep Neural Network for Semi-supervised Semantic Segmentation, arXiv:1506.04924. http://arxiv.org/pdf/1506.04924
9. Seunghoon Hong, Junhyuk Oh, Bohyung Han, and Honglak Lee, Learning Transferrable Knowledge for Semantic Segmentation with Deep Convolutional Neural Network. http://arxiv.org/pdf/1512.07928.pdf
Project Pagehttp://cvlab.postech.ac.kr/research/transfernet/
10. Shuai Zheng, Sadeep Jayasumana, Bernardino Romera-Paredes, Vibhav Vineet, Zhizhong Su, Dalong Du, Chang Huang, Philip H. S. Torr, Conditional Random Fields as Recurrent Neural Networks.
http://arxiv.org/pdf/1502.03240
11. Liang-Chieh Chen, George Papandreou, Kevin Murphy, Alan L. Yuille, Weakly-and semi-supervised learning of a DCNN for semantic image segmentation, arXiv:1502.02734. http://arxiv.org/pdf/1502.02734
12. Mohammadreza Mostajabi, Payman Yadollahpour, Gregory Shakhnarovich, Feedforward Semantic Segmentation With Zoom-Out Features, CVPR, 2015 http://www.cv-foundation.org/openaccess/content*cvpr*2015/papers/Mostajabi*Feedforward*Semantic*Segmentation*2015*CVPR*paper.pdf
13. Holger Caesar, Jasper Uijlings, Vittorio Ferrari, Joint Calibration for Semantic Segmentation. http://arxiv.org/pdf/1507.01581
14. Jonathan Long, Evan Shelhamer, Trevor Darrell, Fully Convolutional Networks for Semantic Segmentation, CVPR, 2015.
http://www.cv-foundation.org/openaccess/content*cvpr*2015/papers/Long*Fully*Convolutional*Networks*2015*CVPR*paper.pdf
15. Bharath Hariharan, Pablo Arbelaez, Ross Girshick, Jitendra Malik, Hypercolumns for Object Segmentation and Fine-Grained Localization, CVPR, 2015.

http://www.cv-foundation.org/openaccess/content*cvpr*2015/papers/Hariharan*Hyper-columns*for*Object*2015*CVPR*paper.pdf

16. Abhishek Sharma, Oncel Tuzel, David W. Jacobs, Deep Hierarchical Parsing for Semantic Segmentation, CVPR, 2015. http://www.cv-foundation.org/openaccess/content*cvpr*2015/papers/Sharma*Deep*Hierarchical*Parsing*2015*CVPR*paper.pdf
17. Clement Farabet, Camille Couprie, Laurent Najman, Yann LeCun, Scene Parsing with Multiscale Feature Learning, Purity Trees, and Optimal Covers, ICML, 2012. http://yann.lecun.com/exdb/publis/pdf/farabet-icml-12.pdf
18. Clement Farabet, Camille Couprie, Laurent Najman, Yann LeCun, Learning Hierarchical Features for Scene Labeling, PAMI, 2013. http://yann.lecun.com/exdb/publis/pdf/farabet-pami-13.pdf
19. Fisher Yu, Vladlen Koltun, "Multi-Scale Context Aggregation by Dilated Convolutions", ICLR 2016, http://arxiv.org/pdf/1511.07122v2.pdf
20. Niloufar Pourian, S. Karthikeyan, and B.S. Manjunath, "Weakly supervised graph based semantic segmentation by learning communities of image-parts", ICCV, 2015, http://www.cv-foundation.org/openaccess/content*iccv*2015/papers/Pourian*Weakly*Supervised*Graph*ICCV*2015*paper.pdf

Object Recognition

1. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, Trevor Darrell http://arxiv.org/abs/1310.1531
2. CNN Features off-the-shelf: an Astounding Baseline for Recognition CVPR 2014 http://arxiv.org/abs/1403.6382

3. HD-CNN: Hierarchical Deep Convolutional Neural Network for Image Classification intro: ICCV 2015
https://arxiv.org/abs/1410.0736
4. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. ImageNet top-5 error: 4.94%
http://arxiv.org/abs/1502.01852
5. Humans and deep networks largely agree on which kinds of variation make object recognition harder
http://arxiv.org/abs/1604.06486
6. FusionNet: 3D Object Classification Using Multiple Data Representations
https://arxiv.org/abs/1607.05695
7. Deep FisherNet for Object Classification
http://arxiv.org/abs/1608.00182
8. Factorized Bilinear Models for Image Recognition
https://arxiv.org/abs/1611.05709
9. Hyperspectral CNN Classification with Limited Training Samples
https://arxiv.org/abs/1611.09007
10. The More You Know: Using Knowledge Graphs for Image Classification
https://arxiv.org/abs/1612.04844
11. MaxMin Convolutional Neural Networks for Image Classification
http://webia.lip6.fr/
12. Cost-Effective Active Learning for Deep Image Classification. TCSVT 2016.
https://arxiv.org/abs/1701.03551
13. Deep Collaborative Learning for Visual Recognition
https://www.arxiv.org/abs/1703.01229
14. Convolutional Low-Resolution Fine-Grained Classification
https://arxiv.org/abs/1703.05393
15. Deep Mixture of Diverse Experts for Large-Scale Visual Recognition

https://arxiv.org/abs/1706.07901

Sunrise or Sunset: Selective Comparison Learning for Subtle Attribute Recognition

https://arxiv.org/abs/1707.06335

16. Why Do Deep Neural Networks Still Not Recognize These Images?: A Qualitative Analysis on Failure Cases of ImageNet Classification

https://arxiv.org/abs/1709.03439

17. B-CNN: Branch Convolutional Neural Network for Hierarchical Classification

https://arxiv.org/abs/1709.09890(

18. Multiple Object Recognition with Visual Attention

https://arxiv.org/abs/1412.7755

19. Multiple Instance Learning Convolutional Neural Networks for Object Recognition

https://arxiv.org/abs/1610.03155

20. Deep Learning Face Representation from Predicting 10,000 Classes. intro: CVPR 2014

http://mmlab.ie.cuhk.edu.hk/pdf/YiSun*CVPR14.pdf*

21. Deep Learning Face Representation by Joint Identification-Verification

http://papers.nips.cc/paper/5416-analog-memories-in-a-balanced-rate-based-network-of-e-i-neurons

22. Deeply learned face representations are sparse, selective, and robust

http://arxiv.org/abs/1412.1265

23. FaceNet: A Unified Embedding for Face Recognition and Clustering

http://arxiv.org/abs/1503.03832

24. Bilinear CNN Models for Fine-grained Visual Recognition

http://vis-www.cs.umass.edu/bcnn/

25. DeepFood: Deep Learning-Based Food Image Recognition for Computer-Aided Dietary Assessment

http://arxiv.org/abs/1606.05675

26. Multi-attribute Learning for Pedestrian Attribute Recognition in Surveillance Scenarios

http://or.nsfc.gov.cn/bitstream/00001903-5/417802/1/1000014103914.pdf

Image Captioning

1. m-RNN 模型《 Explain Images with Multimodal Recurrent Neural Networks》 2014
   https://arxiv.org/pdf/1410.1090.pdf
2. NIC 模型 《Show and Tell: A Neural Image Caption Generator》2014
3. MS Captivator From captions to visual concepts and back 2014
4. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention 2015
5. What Value Do Explicit High Level Concepts Have in Vision to Language Problems? 2016
   https://arxiv.org/pdf/1506.01144.pdf
6. Guiding Long-Short Term Memory for Image Caption Generation 2015
   https://arxiv.org/pdf/1509.04942.pdf
7. Watch What You Just Said: Image Captioning with Text-Conditional Attention 2016
   https://arxiv.org/pdf/1606.04621.pdf
   https://github.com/LuoweiZhou/e2e-gLSTM-sc
8. Generating Natural-Language Video Descriptions Using Text-Mined Knowledge,2014
   https://www.aaai.org/ocs/index.php/AAAI/AAAI13/paper/view/6454/7204
9. Self-critical Sequence Training for Image Captioning 2017 CVPR
   https://arxiv.org/pdf/1612.00563.pdf
10. Deep Reinforcement Learning-based Image Captioning with Embedding Reward 2017 cvpr
    https://arxiv.org/abs/1704.03899
11. Knowing When to Look: Adaptive Attention via a Visual Sentinel for Image Captioning 2017 cvpr
    https://arxiv.org/pdf/1612.01887.pdf
    https://github.com/jiasenlu/AdaptiveAttention
12. Ryan Kiros, Ruslan Salakhutdinov, Richard S. Zemel, Unifying Visual-Semantic Embeddings with Multimodal Neural Language Models, arXiv:1411.2539.
    https://arxiv.org/abs/1411.2539

13. Berkeley Jeff Donahue, Lisa Anne Hendricks, Sergio Guadarrama, Marcus Rohrbach, Subhashini Venugopalan, Kate Saenko, Trevor Darrell, Long-term Recurrent Convolutional Networks for Visual Recognition and Description
https://arxiv.org/abs/1411.4389
14. UML / UT Subhashini Venugopalan, Huijuan Xu, Jeff Donahue, Marcus Rohrbach, Raymond Mooney, Kate Saenko, Translating Videos to Natural Language Using Deep Recurrent Neural Networks, NAACL-HLT, 2015.
https://arxiv.org/abs/1412.4729
15. CMU / Microsoft Xinlei Chen, C. Lawrence Zitnick, Learning a Recurrent Visual Representation for Image Caption Generation
https://arxiv.org/abs/1411.5654
16. Xinlei Chen, C. Lawrence Zitnick, Mind's Eye: A Recurrent Visual Representation for Image Caption Generation, CVPR 2015
https://www.cs.cmu.edu/~xinleic/papers/cvpr15_rnn.pdf
17. Facebook Remi Lebret, Pedro O. Pinheiro, Ronan Collobert, Phrase-based Image Captioning, arXiv:1502.03671 / ICML 2015
https://arxiv.org/abs/1502.03671
18. UCLA / Baidu Junhua Mao, Wei Xu, Yi Yang, Jiang Wang, Zhiheng Huang, Alan L. Yuille, Learning like a Child: Fast Novel Visual Concept Learning from Sentence Descriptions of Images
https://arxiv.org/abs/1504.06692
19. MS + Berkeley Jacob Devlin, Saurabh Gupta, Ross Girshick, Margaret Mitchell, C. Lawrence Zitnick, Exploring Nearest Neighbor Approaches for Image Captioning
https://arxiv.org/abs/1505.04467
20. Jacob Devlin, Hao Cheng, Hao Fang, Saurabh Gupta, Li Deng, Xiaodong He, Geoffrey Zweig, Margaret Mitchell, Language Models for Image Captioning: The Quirks and What Works
https://arxiv.org/abs/1505.01809

21. Adelaide Qi Wu, Chunhua Shen, Anton van den Hengel, Lingqiao Liu, Anthony Dick, Image Captioning with an Intermediate Attributes Layer
https://arxiv.org/abs/1506.01144v1
22. Tilburg Grzegorz Chrupala, Akos Kadar, Afra Alishahi, Learning language through pictures
https://arxiv.org/abs/1506.03694
23. Univ. Montreal Kyunghyun Cho, Aaron Courville, Yoshua Bengio, Describing Multi-media Content using Attention-based Encoder-Decoder Networks
https://arxiv.org/abs/1507.01053
24. Cornell Jack Hessel, Nicolas Savva, Michael J. Wilber, Image Representations and New Domains in Neural Image Captioning
https://arxiv.org/abs/1508.02091
25. MS + City Univ. of HongKong Ting Yao, Tao Mei, and Chong-Wah Ngo, "Learning Query and Image Similarities with Ranking Canonical Correlation Analysis", ICCV, 2015
https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Yao_Learning_Query_and_ICCV_2015_paper.pdf
26. Mao J, Xu W, Yang Y, et al. Deep Captioning with Multimodal Recurrent Neural Networks (m- RNN) 2015.
https://arxiv.org/abs/1412.6632
27. Pan Y, Yao T, Li H, et al. Video Captioning with Transferred Semantic Attributes 2016.
https://arxiv.org/abs/1611.07675
28. Johnson J, Karpathy A, Li F F. DenseCap: Fully Convolutional Localization Networks for Dense Captioning
https://arxiv.org/abs/1511.07571
29. I. Sutskever, O. Vinyals, and Q. V. Le. Sequence to sequence learning with neural networks. NIPS 2014.
https://arxiv.org/abs/1409.3215

30. Karpathy A, Li F F. Deep Visual-Semantic Alignments for Generating Image Descriptions TPAMI 2015
https://arxiv.org/abs/1412.2306
31. A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei. Large-scale video classification with convolutional neural networks. CVPR, 2014.
http://cs.stanford.edu/people/karpathy/deepvideo/
32. Yao T, Pan Y, Li Y, et al. Boosting Image Captioning with Attributes 2016.
https://arxiv.org/abs/1611.01646
33. Venugopalan S, Rohrbach M, Donahue J, et al. Sequence to Sequence -- Video to Text. 2015.
https://arxiv.org/abs/1505.00487
34. Captioning Images with Diverse Objects, CVPR 2017
35. Dense Captioning with Joint Inference and Visual Context, CVPR 2017
36. Incorporating Copying Mechanism in Image Captioning, CVPR 2017
37. Skeleton Key Image Captioning by Skeleton-Attribute Decomposition, CVPR 2017
38. An Empirical Study of Language CNN for Image Captioning, ICCV 2017
39. Areas of Attention for Image Captioning, ICCV 2017
40. Improved Image Captioning via Policy Gradient optimization of SPIDEr, ICCV 2017
41. Paying Attention to Descriptions Generated by Image Captioning Models, ICCV 2017
42. Scene Graph Generation from Objects, Phrases and Region Captions, ICCV 2017
43. Show, Adapt and Tell Adversarial Training of Cross-domain Image Captioner, ICCV 2017
44. Speaking the Same Language Matching Machine to Human Captions by Adversarial Training, ICCV 2017
45. Paying Attention to Descriptions Generated by Image Captioning Models, ICCV 2017

Video Captioning

. Jeff Donahue, Lisa Anne Hendricks, Sergio Guadarrama, Marcus Rohrbach, Subhashini Venugopalan, Kate Saenko, Trevor Darrell, Long-term Recurrent Convolutional Networks for Visual Recognition and Description, CVPR, 2015.

http://jeffdonahue.com/lrcn/

http://arxiv.org/pdf/1411.4389.pdf

1. Subhashini Venugopalan, Huijuan Xu, Jeff Donahue, Marcus Rohrbach, Raymond Mooney, Kate Saenko, Translating Videos to Natural Language Using Deep Recurrent Neural Networks, arXiv:1412.4729.
   UT / UML / Berkeley http://arxiv.org/pdf/1412.4729
2. Yingwei Pan, Tao Mei, Ting Yao, Houqiang Li, Yong Rui, Joint Modeling Embedding and Translation to Bridge Video and Language, arXiv:1505.01861.
   Microsoft http://arxiv.org/pdf/1505.01861
3. Subhashini Venugopalan, Marcus Rohrbach, Jeff Donahue, Raymond Mooney, Trevor Darrell, Kate Saenko, Sequence to Sequence--Video to Text, arXiv:1505.00487.
   UT / Berkeley / UML http://arxiv.org/pdf/1505.00487
4. Li Yao, Atousa Torabi, Kyunghyun Cho, Nicolas Ballas, Christopher Pal, Hugo Larochelle, Aaron Courville, Describing Videos by Exploiting Temporal Structure, arXiv:1502.08029
   Univ. Montreal / Univ. Sherbrooke http://arxiv.org/pdf/1502.08029.pdf]
5. Anna Rohrbach, Marcus Rohrbach, Bernt Schiele, The Long-Short Story of Movie Description, arXiv:1506.01698
   MPI / Berkeley http://arxiv.org/pdf/1506.01698.pdf]
6. Yukun Zhu, Ryan Kiros, Richard Zemel, Ruslan Salakhutdinov, Raquel Urtasun, Antonio Torralba, Sanja Fidler, Aligning Books and Movies: Towards Story-like Visual Explanations by Watching Movies and Reading Books, arXiv:1506.06724
   Univ. Toronto / MIT [http://arxiv.org/pdf/1506.06724.pdf]

7. Kyunghyun Cho, Aaron Courville, Yoshua Bengio, Describing Multimedia Content using Attention-based Encoder-Decoder Networks, arXiv:1507.01053
Univ. Montreal http://arxiv.org/pdf/1507.01053.pdf
8. Dotan Kaufman, Gil Levi, Tal Hassner, Lior Wolf, Temporal Tessellation for Video Annotation and Summarization, arXiv:1612.06950.
TAU / USC [https://arxiv.org/pdf/1612.06950.pdf]
9. Chiori Hori, Takaaki Hori, Teng-Yok Lee, Kazuhiro Sumi, John R. Hershey, Tim K. Marks Attention-Based Multimodal Fusion for Video Description
https://arxiv.org/abs/1701.03126
10. Describing Videos using Multi-modal Fusion
https://dl.acm.org/citation.cfm?id=2984065
11. Andrew Shin , Katsunori Ohnishi , Tatsuya Harada Beyond caption to narrative: Video captioning with multiple sentences
http://ieeexplore.ieee.org/abstract/document/7532983/
12. Jianfeng Dong, Xirong Li, Cees G. M. Snoek Word2VisualVec: Image and Video to Sentence Matching by Visual Feature Prediction
https://pdfs.semanticscholar.org/de22/8875bc33e9db85123469ef80fc0071a92386.pdf
13. Multimodal Video Description
https://dl.acm.org/citation.cfm?id=2984066
14. Xiaodan Liang, Zhiting Hu, Hao Zhang, Chuang Gan, Eric P. Xing Recurrent Topic-Transition GAN for Visual Paragraph Generation
https://arxiv.org/abs/1703.07022
15. Weakly Supervised Dense Video Captioning（CVPR2017）
16. Multi-Task Video Captioning with Video and Entailment Generation（ACL2017）

Visual Question Answering

1. Kushal Kafle, and Christopher Kanan. *An Analysis of Visual Question Answering Algorithms.* arXiv:1703.09684, 2017. https://arxiv.org/abs/1703.09684

2. Hyeonseob Nam, Jung-Woo Ha, Jeonghee Kim, Dual Attention Networks for Multi-modal Reasoning and Matching, arXiv:1611.00471, 2016. https://arxiv.org/abs/1611.00471
3. Jin-Hwa Kim, Kyoung Woon On, Jeonghee Kim, Jung-Woo Ha, Byoung-Tak Zhang, Hadamard Product for Low-rank Bilinear Pooling, arXiv:1610.04325, 2016. https://arxiv.org/abs/1610.04325
4. Akira Fukui, Dong Huk Park, Daylen Yang, Anna Rohrbach, Trevor Darrell, Marcus Rohrbach, Multimodal Compact Bilinear Pooling for Visual Question Answering and Visual Grounding, arXiv:1606.01847, 2016. https://arxiv.org/abs/1606.01847 [code]https://github.com/akirafukui/vqa-mcb
5. Kuniaki Saito, Andrew Shin, Yoshitaka Ushiku, Tatsuya Harada, DualNet: Domain-Invariant Network for Visual Question Answering. arXiv:1606.06108v1, 2016. https://arxiv.org/pdf/1606.06108.pdf
6. Arijit Ray, Gordon Christie, Mohit Bansal, Dhruv Batra, Devi Parikh, Question Relevance in VQA: Identifying Non-Visual And False-Premise Questions, arXiv:1606.06622, 2016. https://arxiv.org/pdf/1606.06622v1.pdf
7. Hyeonwoo Noh, Bohyung Han, Training Recurrent Answering Units with Joint Loss Minimization for VQA, arXiv:1606.03647, 2016. http://arxiv.org/abs/1606.03647v1
8. Jiasen Lu, Jianwei Yang, Dhruv Batra, Devi Parikh, Hierarchical Question-Image Co-Attention for Visual Question Answering, arXiv:1606.00061, 2016. https://arxiv.org/pdf/1606.00061v2.pdf [code]https://github.com/jiasenlu/HieCoAtten-VQA
9. Jin-Hwa Kim, Sang-Woo Lee, Dong-Hyun Kwak, Min-Oh Heo, Jeonghee Kim, Jung-Woo Ha, Byoung-Tak Zhang, Multimodal Residual Learning for Visual QA, arXiv:1606.01455, 2016. https://arxiv.org/pdf/1606.01455v1.pdf
10. Peng Wang, Qi Wu, Chunhua Shen, Anton van den Hengel, Anthony Dick, FVQA: Fact-based Visual Question Answering, arXiv:1606.05433, 2016. https://arxiv.org/pdf/1606.05433.pdf

11. Ilija Ilievski, Shuicheng Yan, Jiashi Feng, A Focused Dynamic Attention Model for Visual Question Answering, arXiv:1604.01485. https://arxiv.org/pdf/1604.01485v1.pdf
12. Yuke Zhu, Oliver Groth, Michael Bernstein, Li Fei-Fei, Visual7W: Grounded Question Answering in Images, CVPR 2016. http://arxiv.org/abs/1511.03416
13. Hyeonwoo Noh, Paul Hongsuck Seo, and Bohyung Han, Image Question Answering using Convolutional Neural Network with Dynamic Parameter Prediction, CVPR, 2016.http://arxiv.org/pdf/1511.05756.pdf
14. Jacob Andreas, Marcus Rohrbach, Trevor Darrell, Dan Klein, Learning to Compose Neural Networks for Question Answering, NAACL 2016. http://arxiv.org/pdf/1601.01705.pdf
15. Jacob Andreas, Marcus Rohrbach, Trevor Darrell, Dan Klein, Deep compositional question answering with neural module networks, CVPR 2016. https://arxiv.org/abs/1511.02799
16. Zichao Yang, Xiaodong He, Jianfeng Gao, Li Deng, Alex Smola, Stacked Attention Networks for Image Question Answering, CVPR 2016. http://arxiv.org/abs/1511.02274 [code]https://github.com/JamesChuanggg/san-torch
17. Kevin J. Shih, Saurabh Singh, Derek Hoiem, Where To Look: Focus Regions for Visual Question Answering, CVPR, 2015. http://arxiv.org/pdf/1511.07394v2.pdf
18. Kan Chen, Jiang Wang, Liang-Chieh Chen, Haoyuan Gao, Wei Xu, Ram Nevatia, ABC-CNN: An Attention Based Convolutional Neural Network for Visual Question Answering, arXiv:1511.05960v1, Nov 2015. http://arxiv.org/pdf/1511.05960v1.pdf
19. Huijuan Xu, Kate Saenko, Ask, Attend and Answer: Exploring Question-Guided Spatial Attention for Visual Question Answering, arXiv:1511.05234v1, Nov 2015. http://arxiv.org/abs/1511.05234
20. Kushal Kafle and Christopher Kanan, Answer-Type Prediction for Visual Question Answering, CVPR 2016. http://www.cv-foundation.org/openaccess/content*cvpr*2016/html/Kafle*Answer-Type*Prediction*for*CVPR*2016*paper.html

21. Stanislaw Antol, Aishwarya Agrawal, Jiasen Lu, Margaret Mitchell, Dhruv Batra, C. Lawrence Zitnick, Devi Parikh, VQA: Visual Question Answering, ICCV, 2015. http://arxiv.org/pdf/1505.00468
22. Stanislaw Antol, Aishwarya Agrawal, Jiasen Lu, Margaret Mitchell, Dhruv Batra, C. Lawrence Zitnick, Devi Parikh, VQA: Visual Question Answering, ICCV, 2015. http://arxiv.org/pdf/1505.00468 [code]https://github.com/JamesChuanggg/VQA-tensorflow
23. Bolei Zhou, Yuandong Tian, Sainbayar Sukhbaatar, Arthur Szlam, Rob Fergus, Simple Baseline for Visual Question Answering, arXiv:1512.02167v2, Dec 2015. http://arxiv.org/abs/1512.02167
24. Hauyuan Gao, Junhua Mao, Jie Zhou, Zhiheng Huang, Lei Wang, Wei Xu, Are You Talking to a Machine? Dataset and Methods for Multilingual Image Question Answering, NIPS 2015. http://arxiv.org/pdf/1505.05612.pdf
25. Mateusz Malinowski, Marcus Rohrbach, Mario Fritz, Ask Your Neurons: A Neural-based Approach to Answering Questions about Images, ICCV 2015. http://arxiv.org/pdf/1505.01121v3.pdf
26. Mengye Ren, Ryan Kiros, Richard Zemel, Exploring Models and Data for Image Question Answering, ICML 2015. http://arxiv.org/pdf/1505.02074.pdf
27. Mateusz Malinowski, Mario Fritz, Towards a Visual Turing Challe, NIPS Workshop 2015. http://arxiv.org/abs/1410.8027
28. Mateusz Malinowski, Mario Fritz, A Multi-World Approach to Question Answering about Real-World Scenes based on Uncertain Input, NIPS 2014. http://arxiv.org/pdf/1410.0210v4.pdf
29. Hedi Ben-younes, Remi Cadene, Matthieu Cord, Nicolas Thome: MUTAN: Multimodal Tucker Fusion for Visual Question Answering https://arxiv.org/pdf/1705.06676.pdf [Code]https://github.com/Cadene/vqa.pytorch

30. Jin-Hwa Kim, Kyoung Woon On, Jeonghee Kim, Jung-Woo Ha, Byoung-Tak Zhang, Hadamard Product for Low-rank Bilinear Pooling, arXiv:1610.04325, 2016. https://arxiv.org/abs/1610.04325
31. Akira Fukui, Dong Huk Park, Daylen Yang, Anna Rohrbach, Trevor Darrell, Marcus Rohrbach, Multimodal Compact Bilinear Pooling for Visual Question Answering and Visual Grounding, arXiv:1606.01847, 2016. https://arxiv.org/abs/1606.01847

Edge Detection

1. Saining Xie, Zhuowen Tu, Holistically-Nested Edge Detection Holistically-Nested Edge Detection
http://arxiv.org/pdf/1504.06375
https://github.com/s9xie/hed
2. Gedas Bertasius, Jianbo Shi, Lorenzo Torresani, DeepEdge: A Multi-Scale Bifurcated Deep Network for Top-Down Contour Detection, CVPR, 2015.
http://arxiv.org/pdf/1412.1123
3. Wei Shen, Xinggang Wang, Yan Wang, Xiang Bai, Zhijiang Zhang, DeepContour: A Deep Convolutional Feature Learned by Positive-Sharing Loss for Contour Detection, CVPR, 2015.
http://mc.eistar.net/UpLoadFiles/Papers/DeepContour*cvpr15.pdf*

Human Pose Estimation

1. Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh, Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, CVPR, 2017.
2. Leonid Pishchulin, Eldar Insafutdinov, Siyu Tang, Bjoern Andres, Mykhaylo Andriluka, Peter Gehler, and Bernt Schiele, Deepcut: Joint subset partition and labeling for multi person pose estimation, CVPR, 2016.

3. Shih-En Wei, Varun Ramakrishna, Takeo Kanade, and Yaser Sheikh, Convolutional pose machines, CVPR, 2016.
4. Alejandro Newell, Kaiyu Yang, and Jia Deng, Stacked hourglass networks for human pose estimation, ECCV, 2016.
5. Tomas Pfister, James Charles, and Andrew Zisserman, Flowing convnets for human pose estimation in videos, ICCV, 2015.
6. Jonathan J. Tompson, Arjun Jain, Yann LeCun, Christoph Bregler, Joint training of a convolutional network and a graphical model for human pose estimation, NIPS, 2014.

Image Generation

1. Aäron van den Oord, Nal Kalchbrenner, Oriol Vinyals, Lasse Espeholt, Alex Graves, Koray Kavukcuoglu. "Conditional Image Generation with PixelCNN Decoders"https://arxiv.org/pdf/1606.05328v2.pdfhttps://github.com/kundan2510/pixelCNN
2. Alexey Dosovitskiy, Jost Tobias Springenberg, Thomas Brox, "Learning to Generate Chairs with Convolutional Neural Networks", CVPR, 2015. http://www.cv-foundation.org/openaccess/content*cvpr*2015/papers/Dosovitskiy*Learning*to*Generate*2015*CVPR*paper.pdf
3. Karol Gregor, Ivo Danihelka, Alex Graves, Danilo Jimenez Rezende, Daan Wierstra, "DRAW: A Recurrent Neural Network For Image Generation", ICML, 2015. https://arxiv.org/pdf/1502.04623v2.pdf
4. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, Generative Adversarial Networks, NIPS, 2014. http://arxiv.org/abs/1406.2661
5. Emily Denton, Soumith Chintala, Arthur Szlam, Rob Fergus, Deep Generative Image Models using a ￼Laplacian Pyramid of Adversarial Networks, NIPS, 2015. http://arxiv.org/abs/1506.05751
6. Lucas Theis, Aäron van den Oord, Matthias Bethge, "A note on the evaluation of generative models", ICLR 2016. http://arxiv.org/abs/1511.01844

7. Zhenwen Dai, Andreas Damianou, Javier Gonzalez, Neil Lawrence, "Variationally Auto-Encoded Deep Gaussian Processes", ICLR 2016. http://arxiv.org/pdf/1511.06455v2.pdf
8. Elman Mansimov, Emilio Parisotto, Jimmy Ba, Ruslan Salakhutdinov, "Generating Images from Captions with Attention", ICLR 2016, http://arxiv.org/pdf/1511.02793v2.pdf
9. Jost Tobias Springenberg, "Unsupervised and Semi-supervised Learning with Categorical Generative Adversarial Networks", ICLR 2016, http://arxiv.org/pdf/1511.06390v1.pdf
10. Harrison Edwards, Amos Storkey, "Censoring Representations with an Adversary", ICLR 2016, http://arxiv.org/pdf/1511.05897v3.pdf
11. Takeru Miyato, Shin-ichi Maeda, Masanori Koyama, Ken Nakae, Shin Ishii, "Distributional Smoothing with Virtual Adversarial Training", ICLR 2016, http://arxiv.org/pdf/1507.00677v8.pdf
12. Jun-Yan Zhu, Philipp Krahenbuhl, Eli Shechtman, and Alexei A. Efros, "Generative Visual Manipulation on the Natural Image Manifold", ECCV 2016. https://arxiv.org/pdf/1609.03552v2.pdf https://github.com/junyanz/iGAN https://youtu.be/9c4z6YsBGQ0
13. Alec Radford, Luke Metz, Soumith Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks", ICLR 2016. http://arxiv.org/pdf/1511.06434.pdf

课程

1. 斯坦福视觉实验室主页：http://vision.stanford.edu/
   李飞飞组 CS131, CS231A, CS231n 三个课程，可是说是最好的计算机视觉课程。
2. CS 131 Computer Vision: Foundations and Applications：

基础知识：主要讲传统的边缘检测，特征点描述，相机标定，全景图拼接等知识

http://vision.stanford.edu/teaching/cs131*fall1415/schedule.html*

3. CS231A Computer Vision: from 3D reconstruction to recognition:

   http://cvgl.stanford.edu/teaching/cs231a*winter1415/schedule.html*

4. CS231n 2017: Convolutional Neural Networks for Visual Recognition

   主要讲卷积神经网络的具体结构，各组成部分的原理优化以及各种应用。

   http://vision.stanford.edu/teaching/cs231n/

   国内地址：http://www.bilibili.com/video/av13260183/

5. Stanford CS231n 2016 : Convolutional Neural Networks for Visual Recognition
   - homepage: http://cs231n.stanford.edu/
   - homepage: http://vision.stanford.edu/teaching/cs231n/index.html
   - syllabus: http://vision.stanford.edu/teaching/cs231n/syllabus.html
   - course notes: http://cs231n.github.io/
   - youtube: https://www.youtube.com/watch?v=NfnWJUyUJYU&feature=youtu.be
   - mirror: http://pan.baidu.com/s/1pKsTivp
   - mirror: http://pan.baidu.com/s/1c2wR8dy
   - 网易中文字幕：http://study.163.com/course/introduction/1003223001.htm
   - assignment 1: http://cs231n.github.io/assignments2016/assignment1/
   - assignment 2: http://cs231n.github.io/assignments2016/assignment2/
   - assignment 3: http://cs231n.github.io/assignments2016/assignment3/
6. 1st Summer School on Deep Learning for Computer Vision Barcelona: (July 4-8, 2016)
   - 深度学习计算机视觉夏季学校课程， 包含基础知识以及许多深度学习在计算机视觉中的应用，比如分类，检测，captioning 等等
   - homepage(slides+videos): http://imatge-upc.github.io/telecombcn-2016-dlcv/
   - homepage: https://imatge.upc.edu/web/teaching/deep-learning-computer-vision
   - youtube: https://www.youtube.com/user/imatgeupc/videos?shelf*id=0&sort=dd&view=0*

7. 2nd Summer School on Deep Learning for Computer VisionBarcelona (June 21-27, 2017)
   https://telecombcn-dl.github.io/2017-dlcv/

综述

1. Annotated Computer Vision Bibliography: Table of Contents. Since 1994
   Keith Price 从 1994 年开始做了这个索引，涵盖了所有计算机视觉里面所有 topic，所有 subtopic 的著作，包括论文，教材，还对各类主题的关键词。这个网站频繁更新（最近一次是 2017 年 8 月 28 号），收录每个方向重要期刊，会议文献和书籍，并且保证了所有链接不失效。
2. What Sparked Video Research in 1877? The Overlooked Role of the Siemens Artificial Eye by Mark Schubin 2017
   http://ieeexplore.ieee.org/document/7857854/
3. Giving machines humanlike eyes. by Posch, C., Benosman, R., Etienne-Cummings, R. 2015
   http://ieeexplore.ieee.org/document/7335800/
4. Seeing is not enough by Tom GellerOberlin, OH
   https://dl.acm.org/citation.cfm?id=2001276
5. Visual Tracking: An Experimental Survey
   https://dl.acm.org/citation.cfm?id=2693387
6. A survey on object recognition and segmentation techniques
   http://ieeexplore.ieee.org/document/7724975/
7. A Review of Image Recognition with Deep Convolutional Neural Network
   https://link.springer.com/chapter/10.1007/978-3-319-63309-1_7
8. Recent Advance in Content-based Image Retrieval: A Literature Survey. Wengang Zhou, Houqiang Li, and Qi Tian 2017
   https://arxiv.org/pdf/1706.06064.pdf

9. Automatic Description Generation from Images: A Survey of Models, Datasets, and Evaluation Measures 2016
   https://www.jair.org/media/4900/live-4900-9139-jair.pdf

Turorial

1. Intro to Deep Learning for Computer Vision 2016
   http://chaosmail.github.io/deeplearning/2016/10/22/intro-to-deep-learning-for-computer-vision/
2. CVPR 2014 Tutorial on Deep Learning in Computer Vision
   https://sites.google.com/site/deeplearningcvpr2014/
3. CVPR 2015 Applied Deep Learning for Computer Vision with Torch
   https://github.com/soumith/cvpr2015
4. Deep Learning for Computer Vision – Introduction to Convolution Neural Networks
   http://www.analyticsvidhya.com/blog/2016/04/deep-learning-computer-vision-introduction-convolution-neural-networks/
5. A Beginner's Guide To Understanding Convolutional Neural Networks
   https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/
6. CVPR'17 Tutorial Deep Learning for Objects and Scenes by Kaiming He Ross Girshick
   http://deeplearning.csail.mit.edu/
7. CVPR tutorial : Large-Scale Visual Recognition
   http://www.europe.naverlabs.com/Research/Computer-Vision/Highlights/CVPR-tutorial-Large-Scale-Visual-Recognition
8. CVPR’16 Tutorial on Image Tag Assignment, Refinement and Retrieval
   http://www.lambertoballan.net/2016/06/cvpr16-tutorial-image-tag-assignment-refinement-and-retrieval/
9. Tutorial on Answering Questions about Images with Deep Learning

The tutorial was presented at '2nd Summer School on Integrating Vision and Language: Deep Learning' in Malta, 2016

https://arxiv.org/abs/1610.01076

10. “Semantic Segmentation for Scene Understanding: Algorithms and Implementations" tutorial

https://www.youtube.com/watch?v=pQ318oCGJGY

11. A tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the "echo state network" approach

http://minds.jacobs-university.de/sites/default/files/uploads/papers/ESNTutorialRev.pdf

http://deeplearning.cs.cmu.edu/notes/shaoweiwang.pdf

12. Towards Good Practices for Recognition & Detection by Hikvision Research Institute. Supervised Data Augmentation (SDA)

http://image-net.org/challenges/talks/2016/Hikvision*at*ImageNet*2016.pdf*

13. Generative Adversarial Networks by Ian Goodfellow, NIPS 2016 tutorial

https://arxiv.org/abs/1701.00160

http://www.iangoodfellow.com/slides/2016-12-04-NIPS.pdf

14. Deep Learning for Computer Vision – Introduction to Convolution Neural Networks

http://www.analyticsvidhya.com/blog/2016/04/deep-learning-computer-vision-introduction-convolution-neural-networks/

图书

1. 两本经典教材《Computer Vision: A Modern Approach》和《Computer Vision: Algorithms and Applications》，可以先读完第一本再读第二本。
2. Computer Vision: A Modern Approach by David A. Forsyth, Jean Ponce

英文：http://cmuems.com/excap/readings/forsyth-ponce-computer-vision-a-modern-approach.pdf

中文：https://pan.baidu.com/s/1min99eK

3. Computer Vision: Algorithms and Applications by Richard Szeliski

   英文：http://szeliski.org/Book/drafts/SzeliskiBook_20100903_draft.pdf

   中文：https://pan.baidu.com/s/1mhYGtio

4. Computer Vision: Models, Learning, and Inference by Simon J.D. Prince

   书的主页上还有配套的 Slider, 代码，tutorial，演示等各种资源。

   http://www.computervisionmodels.com/

相关期刊与会议

国际会议

1. CVPR, Computer Vision and Pattern Recognition

   CVPR 2017：http://cvpr2017.thecvf.com/

2. ICCV, International Conference on Computer Vision

   ICCV2017：http://iccv2017.thecvf.com/

3. ECCV, European Conference on Computer Vision

4. SIGGRAPH, Special Interest Group on Computer Graphics and Interactive techniques

   SIGGRAPH2017 http://s2017.siggraph.org/

5. ACM International Conference on Multimedia

   ACMMM2017： http://www.acmmm.org/2017/

6. ICIP, International Conference on Image Processing

   http://2017.ieeeicip.org/

期刊

1. ACM Transactions on Graphics, TOG
2. International Journal of Computer Vision, IJCV

3. IEEE Trans on Pattern Analysis and Machine Intelligence, TPAMI
4. IEEE Transactions on Image Processing, TIP
5. IEEE Transactions on Visualization and Computer Graphics, TVCG
6. IEEE Communications Surveys and Tutorials
7. IEEE Signal Processing Magazine
8. IEEE Transactions on EVOLUTIONARY COMPUTATION
9. IEEE Transactions on GEOSCIENCE and REMOTE SENSING 2 区
10. IEEE Transactions on Pattern Analysis and Machine Intelligence
11. NEUROCOMPUTING 2 区
12. Pattern Recognition Letters 2 区
13. Proceedings of the IEEE
14. Signal image and Video Processing 4 区
15. IEEE journal on Selected areas in Communications 2 区
16. IEEE Transactions on image Processing 2 区
17. journal of Visual Communication and image Representation 3 区
18. Machine Vision and Application 3 区
19. Pattern Recognition 2 区
20. Signal Processing-image Communication 3 区
21. COMPUTER Vision and image UNDERSTANDING 3 区
22. IEEE Communications Surveys and Tutorials
23. IET image Processing 4 区
24. Artificial Intelligence 2 区
25. Machine Learning 3 区
26. Medical image Analysis 2 区

领域专家

（水平有限，漏了很多大牛，欢迎大家提建议和补充，会一直保持更新）

华人机构和学者

- 旷视首席科学家, 前 MSRA 首席研究员 孙剑 http://www.jiansun.org/
- 微软微软全球执行副总裁 沈向阳 https://news.microsoft.com/exec/harry-shum/
- 微软亚洲研究院 华刚 https://www.microsoft.com/en-us/research/people/ganghua/
- 上海科技大学的 虞晶怡 http://www.yu-jingyi.com/
- 微软亚洲研究院 梅涛 https://www.microsoft.com/en-us/research/people/tmei/
- 微软亚洲研究院 张正友 https://www.microsoft.com/en-us/research/people/zhang/
- 微软研究院 刘自成 http://people.ucas.ac.cn/~xlchen
- 微软亚洲研究院 王井东 https://www.microsoft.com/en-us/research/people/jingdw/
- 原百度研究院院长 林元庆 https://www.linkedin.com/in/yuanqing-lin-8666789/
- 澳大利亚国立大学 李宏东 http://users.cecs.anu.edu.au/~hongdong/
- 加州大学伯克利分校 马毅 http://yima.csl.illinois.edu/
- 密苏里科技大学 尹朝征 http://web.mst.edu/~yinz/
- 美国西北大学 吴郢 http://www.mccormick.northwestern.edu/research-faculty/directory/profiles/wu-ying.html
- 新加坡国立大学 360 颜水成团队 https://www.ece.nus.edu.sg/stfpage/eleyans/
- 新加坡国立大学 冯佳时 https://sites.google.com/site/jshfeng/home
- 香港中文大学教授贾佳亚：http://www.cse.cuhk.edu.hk/香港中文大学多媒体实验室&商汤（汤晓鸥团队）； http://mmlab.ie.cuhk.edu.hk/,
- https://www.ie.cuhk.edu.hk/people/xotang.shtml
- 香港中文大学教授王晓刚； http://www.ee.cuhk.edu.hk/图森首席科学家，香港科技大学王乃岩博士以及其团队 http://www.winsty.net/
- 美国伊利诺斯大学黄煦涛 https://ece.illinois.edu/directory/profile/t-huang1
- 奥尔巴尼大学陈梅 http://www.albany.edu/meichen/
- 宾夕法尼亚州立大学 刘燕西 http://www.cse.psu.edu/~yul11/
- 亮风台联合创始人、首席科学家 凌海滨及其团队 http://www.dabi.temple.edu/~hbling/

- UCLA 教授朱松纯； http://www.stat.ucla.edu/肯塔基大学计算机系 杨睿刚 http://www.vis.uky.edu/~ryang/
- 南洋理工大学 袁浚菘
- 中科院自动化所； http://www.ia.cas.cn/
- 中科院自动化所模式识别国家重点实验室；http://www.nlpr.ia.ac.cn/CN/model/index.shtml
  - 视觉信息处理研究组
    - 机器视觉课题组：http://vision.ia.ac.cn
    - 图像视频组：http://www.nlpr.ia.ac.cn/iva
  - 生物识别与安全研究组（生物识别与安全技术研究中心）：http://www.cbsr.ia.ac.cn
    - 智能感知与计算研究中心：http://www.cripac.ia.ac.cn
    - Li-Group（李子青组）：http://www.cbsr.ia.ac.cn/Li%20Group/index%20CH.asp,
      中科奥森科技有限公司：http://www.authenmetric.com，“中科奥森”
    - 胡卫明组
  - 模式识别基础理论与方法研究组（Pattern Analysis and Learning Group）：http://www.nlpr.ia.ac.cn/pal/
  - 计算医学研究组（脑网络组研究中心）：http://www.brainnetome.org
  - 空天信息研究中心
  - 多媒体计算研究组：http://nlpr-web.ia.ac.cn/mmc/index.html
- 中科院计算所；http://www.ict.ac.cn/
- 中科院计算所智能信息处理重点实验室；http://iip.ict.ac.cn/
  - 视觉信息处理与学习研究组（http://vipl.ict.ac.cn，下设人脸组、手语组、视频组、视觉建模组、情感计算组、视觉场景理解组、多模态生物特征组、多媒体计算与多模态智能组，中科视拓(北京)科技有限公司：http://seetatech.com）
- 前瞻研究实验室

    - 跨媒体计算课题组（http://mcg.ict.ac.cn）
- 信息工程研究所（http://www.cskaoyan.com/thread-205594-1-1.html）
- 多媒体安全与智能分析研究组
    - 刘偲组：http://liusi-group.com
- 美国罗彻斯特大学教授罗杰波:http://www.cs.rochester.edu/u/jluo/
- 北京大学高文教授及其团队：http://www.jdl.ac.cn/htm-gaowen/
- 清华大学章毓晋教授及其团队：http://www.tsinghua.edu.cn/publish/ee/4157/2010/20101217173552339241557/20101217173552339241557.html
- 清华大学朱军，艾海舟，朱文武，鲁继文教授等
    - http://ml.cs.tsinghua.edu.cn/
    - http://media.cs.tsinghua.edu.cn/~ahz/
    - https://baike.baidu.com/item/%E6%9C%B1%E6%96%87%E6%AD%A6/10181070?fr=aladdin
    - http://www.au.tsinghua.edu.cn/publish/au/1714/2016/20160229104943061296929/20160229104943061296929.*html*
- 西安交通大学人工智能与机器人研究所(郑南宁 龚怡宏)：http://www.aiar.xjtu.edu.cn/
  http://gr.xjtu.edu.cn/web/ygong/home
- 天津大学计算机图形图像与可视计算实验室
- 上海交通大学计算机视觉实验室刘允才教授：http://www.visionlab.sjtu.edu.cn/
- https://cvsjtu.wordpress.com/
- 浙江大学：何晓飞，蔡登，宋明黎，李玺，朱建科，潘纲等老师团队
    - 浙江大学图像技术研究与应用（ITRA）团队：http://www.dvzju.com/
- 中国科学技术大学 查正军 http://auto.ustc.edu.cn/teacher_details.php?i=362
- 南京大学 吴建鑫 https://cs.nju.edu.cn/wujx/
- 中山大学：郑伟诗，林倞教授团队
- 南开：程明明教授团队
- 南京审计大学：吴毅教授(tracking)

- 大连理工大学：卢湖川教授(tracking)
- 厦门大学：纪荣嵘和王菡子教授等
- 华中科技大学：白翔教授团队(text detection)
- 北京邮电大学郭军老师组
- 哈工大： 左旺孟老师团队

North America

- 1.Arizona State University - Center for Cognitive Ubiquitous Computing (CUbiC) |: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 2.Arizona State University - Image, Video, and Usability Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 3.Boston University - Image and Video Computing Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 4.Brown University - Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 5.California Institute of Technology - Computational Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 6.Carnegie Mellon University - CI2CV Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 7.Colorado State University - Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 8.Columbia University - Computer Vision Laboratory |: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 9.Cornell University - Computer Vision Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 10.Darthmouth University - Visual Learning Group: https://computervisiononline.com/blog/awesome-computer-vision-groups

- 11.Duke University - Duke Robotics Intelligence and Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 12.Florida State University - The Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 13.George Mason University - Computer Vision and Robotics Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 14.Georgia Institute of Technology - Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 15.Harvard University - Graphics, Vision & Interaction: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 16.Harvard University - Visual Computing Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 17.Harvard University - Visionlab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 18.Indiana University - Computer Vision lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 19.John Hopkins University - Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 20.Louisiana State University - Image Processing & Computer Vision Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 21.Massachusetts Institute of Technology - MIT Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 22.McGill University - Center for Intelligent Machines: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 23.Michigan State University - Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 24.North Carolina State University - Image Analysis and Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups

- 25.Northeastern University - Computational Vision Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 26.Northwestern University - Computational Vision Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 27.Pennsylvania State University - Laboratory for Perception, Action -Cognition: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 28.Princeton University - Princeton Vision and Robotics: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 29.Purdue University - Purdue Robot Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 30.Queen's University - The Robotics and Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 31.Rensselaer Polytechnic Institute - Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 32.Rutgers University - Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 33.San Diego State University - Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 34.Simon Frazer University - Vision and Media Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 35.Stanford University - Stanford Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 36.Stony Brook University - Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 37.Syracuse University - Smart Vision Systems Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 38.University of Alberta - Computer Vision and Robotics Research Group: https://computervisiononline.com/blog/awesome-computer-vision-groups

- 39.University of Buffalo - Computer Vision and Information Visualization: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 40.University of California, Berkeley - Berkeley Vision and Learning Center: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 41.University of California, Irvine - Computational Vision at UC Irvine: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 42.University of California, Los Angeles - UCLA Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 43.University of California, Riverside - Visualization and Intellig-Systems Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 44.University of California, San Diego - UCSD Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 45.University of California, San Diego - Statistical Visual Computing Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 46.University of California, Santa Barbara - Vision Research Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 47.University of California, Santa Cruz - UCSC Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 48.University of Central Florida - Center for Research in Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 49.University of Central Florida - Computational Imaging Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 50.University of Florida - Laboratory for Computer Vision, Graphics -Medical Imaging: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 51.University of Illinois, Chicago - Robotics: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 52.University of Illinois, Chicago - Machine Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups

- 53.University of Illinois, Urbana Champaign - Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 54.University of Illinois, Urbana Champaign - Computer Vision and Robotic Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 55.University of Louisville - Computer Vision and Image Processing Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 56.University of Maryland - Computer Vision Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 57.University of Maryland - Center for Automation Research: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 58.University of Massachusetts, Amherst - Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 59.University of Miami - Underwater Vision and Imaging Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 60.University of Michigan - Michigan Vision and Learning Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 61.University of Minnesota - Digital Technology Center: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 62.University of North Carolina, Chapel Hill - Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 63.University of North Carolina, Charlotte - Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 64.University of Ottawa - VIVA Lab |: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 65.University of Rochester - Center for Visual Science: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 66.University of Southern California - USC Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups

- 67.University of Texas, Austin - UT-Austin Computer Vision Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 68.University of Tennessee - Imaging, Robotics and Intelligent Syst-Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 69.University of Utah - Utah Robotics Center: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 70.University of Virginia - UVA Vision Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 71.University of Washington - Computer Graphics, Vision, Animation and G-Science: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 72.University of Waterloo- Vision and Image Processing Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 73.University of Wisconsin, Madison - Computer Graphics: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 74.York University - Center for Vision Research |: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 75.York University - Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 76.York University - Active and Attentive Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups

Europe

- 1.Aristotle University of Thessaloniki - Artificial Intelligence Information Analysis: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 2.Charles III University of Madrid - Intelligent Systems Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 3.Czech Technical University - Center for Machine Perception: https://computervisiononline.com/blog/awesome-computer-vision-groups

- 4.Delft Universitiy of Technology - Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 5.Delft Universitiy of Technology - Imaging Physics: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 6.Democritus University of Thrace - Laboratory of Robotics and Automation: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 7.Eindhoven University of Technology - Medical Image Analysis Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 8.ETH Zurich - Computer Vision Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 9.EPFL - Computer Vision Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 10.French Institute for Research in Computer Science and Automation (INRIA)- Perception, Cognition and Interaction: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 11.Graz University of Technology - Institute for Computer Graphics and Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 12.Imperial College London - Imperial CV and Learning Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 13.Italian Institute of Technology - PAVIS |: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 14.IT University of Copenhagen - The Image Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 15.KTH Royal Institute of Technology - Computer Vision and Active Perception Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 16.KU Leuven - Processing Speech and Images: https://computervisiononline.com/blog/awesome-computer-vision-groups

- 17.Max Planck Institute for Informatics - Computer Vision and Multimodal Computing: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 18.RWTH Aachen University - Computer Vision Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 19.RWTH Aachen University - Institute of Man Machine Interaction |: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 20.Technical University Munich - Computer Vision Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 21.Trinity College, Dublin - Graphics, Vision and Visualisation: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 22.University of Amsterdam - Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 23.University of Cambridge - Machine Intelligence Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 24.University of Castilla-La Mancha - VISILAB: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 25.University of Catania - Image Processing Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 26.University of Edinburgh - CALVIN |: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 27.University of Leeds - Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 28.University of Lorraine - Lorraine Laboratory for Computer Research and Applications: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 29.University of Lyon - Informatics Laboratory in Image and Information Systems: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 30.University of Milano-Bicocca - Imaging and Vision Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups

- 31.University of Modena and Reggio Emilia - ImageLab |: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 32.University of Oxford - Torr Vision Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 33.University of Oxford - Visual Geometry Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 34.University of Paderborn - GET Lab - Cognitive Systems Engineering: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 35.University of Twente - Computer Vision & Biometrics: https://computervisiononline.com/blog/awesome-computer-vision-groups

Australia

- 1.The University of Adelaide - Australian Centre for Visual Technologies: https://computervisiononline.com/blog/awesome-computer-vision-groups

Asia and Middle East

- 1.Bilkent University - RETINA Vision and Learning Group: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 2.Ben-Gurion University of the Negev, Beer Sheva - The Interdisciplinary Computational Vision Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 3.Hebrew University of Jerusalem - Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 4.Indian Institute of Science - Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 5.Indian Institute of Science - Video Analytics Laboratory (VAL): https://computervisiononline.com/blog/awesome-computer-vision-groups

- 6.Indian Institute of Technology, Madras - Image Processing and Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 7.Nanyang Technological University - Computer Vision Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 8.National University of Singapore - Computer Vision Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 9.Technion Israel Institute of Technology - Vision and Image Sciences Laboratory: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 10.Technion Israel Institute of Technology - Center for Intelligent Systems: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 11.Tel Aviv University Computer Vision: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 12.University of Tehran - Photogrammetry lab: https://computervisiononline.com/blog/awesome-computer-vision-groups
- 13.Weizmann Institute of Science - Computer Vision Lab: https://computervisiononline.com/blog/awesome-computer-vision-groups

Software

1. Caffehttp://caffe.berkeleyvision.org/
2. PyTorch - Tensors and Dynamic neural networks in Python with strong GPU acceleration https://github.com/pytorch/pytorch
3. CNTK - Microsoft Cognitive Toolkithttps://github.com/Microsoft/CNTK
4. Theanohttp://deeplearning.net/software/theano/
5. cuda-convnethttps://code.google.com/p/cuda-convnet2/
6. DeepLearnToolboxhttps://github.com/rasmusbergpalm/DeepLearnToolbox
7. Deepnethttps://github.com/nitishsrivastava/deepnet
8. Deeppyhttps://github.com/andersbll/deeppy

9. JavaNNhttps://github.com/ivan-vasilev/neuralnetworks
10. hebelhttps://github.com/hannes-brt/hebel
11. Mocha.jlhttps://github.com/pluskid/Mocha.jl
12. OpenDLhttps://github.com/guoding83128/OpenDL
13. cuDNNhttps://developer.nvidia.com/cuDNN
14. MGLhttp://melisgl.github.io/mgl-pax-world/mgl-manual.html
15. Knet.jlhttps://github.com/denizyuret/Knet.jl
16. Nvidia DIGITS - a web app based on Caffehttps://github.com/NVIDIA/DIGITS
17. Neon - Python based Deep Learning Frameworkhttps://github.com/NervanaSystems/neon
18. . Keras - Theano based Deep Learning Libraryhttp://keras.io
19. . Chainer - A flexible framework of neural networks for deep learninghttp://chainer.org/
20. RNNLIB - A recurrent neural network libraryhttp://sourceforge.net/p/rnnl/wiki/Home/
21. Brainstorm - Fast, flexible and fun neural networks.https://github.com/IDSIA/brainstorm
22. Tensorflow - Open source software library for numerical computation using data flow graphshttps://github.com/tensorflow/tensorflow
23. DMTK - Microsoft Distributed Machine Learning Tookithttps://github.com/Microsoft/DMTK
24. Scikit Flow - Simplified interface for TensorFlow mimicking Scikit Learn-https://github.com/google/skflow
25. MXnet - Lightweight, Portable, Flexible Distributed/Mobile Deep Learning framework https://github.com/dmlc/mxnet/
26. Apache SINGA - A General Distributed Deep Learning Platformhttp://singa.incubator.apache.org/
27. DSSTNE - Amazon's library for building Deep Learning modelshttps://github.com/amznlabs/amazon-dsstne

28. SyntaxNet - Google's syntactic parser - A TensorFlow dependency library- https://github.com/tensorflow/models/tree/master/syntaxnet
29. mlpack - A scalable Machine Learning libraryhttp://mlpack.org/
30. Paddle - PArallel Distributed Deep LEarning by Baiduhttps://github.com/baidu/paddle
31. NeuPy - Theano based Python library for ANN and Deep Learninghttp://neupy.com
32. Sonnet - a library for constructing neural networks by Google's Deep-Mindhttps://github.com/deepmind/sonnet

Datasets

Detection

1. PASCAL VOC 2009 dataset

   Classification/Detection Competitions, Segmentation Competition, Person Layout Taster Competition datasets
2. LabelMe dataset

   LabelMe is a web-based image annotation tool that allows researchers to label images and share the annotations with the rest of the community. If you use the database, we only ask that you contribute to it, from time to time, by using the labeling tool.
3. BioID Face Detection Database
4. 1521 images with human faces, recorded under natural conditions, i.e. varying illumination and complex background. The eye positions have been set manually.
5. CMU/VASC & PIE Face dataset
6. Yale Face dataset
7. Caltech

   Cars, Motorcycles, Airplanes, Faces, Leaves, Backgrounds
8. Caltech 101

   Pictures of objects belonging to 101 categories

9. Caltech 256

   Pictures of objects belonging to 256 categories

10. Daimler Pedestrian Detection Benchmark

    15,560 pedestrian and non-pedestrian samples (image cut-outs) and 6744 additional full images not containing pedestrians for bootstrapping. The test set contains more than 21,790 images with 56,492 pedestrian labels (fully visible or partially occluded), captured from a vehicle in urban traffic.

11. MIT Pedestrian dataset

    CVC Pedestrian Datasets

12. CVC Pedestrian Datasets

    CBCL Pedestrian Database

13. MIT Face dataset

    CBCL Face Database

14. MIT Car dataset

    CBCL Car Database

15. MIT Street dataset

    CBCL Street Database

16. INRIA Person Data Set

    A large set of marked up images of standing or walking people

17. INRIA car dataset

    A set of car and non-car images taken in a parking lot nearby INRIA

18. INRIA horse dataset

    A set of horse and non-horse images

19. H3D Dataset

    3D skeletons and segmented regions for 1000 people in images

20. HRI RoadTraffic dataset

    A large-scale vehicle detection dataset

21. BelgaLogos

10000 images of natural scenes, with 37 different logos, and 2695 logos instances, annotated with a bounding box.

22. FlickrBelgaLogos

    10000 images of natural scenes grabbed on Flickr, with 2695 logos instances cut and pasted from the BelgaLogos dataset.

23. FlickrLogos-32

    The dataset FlickrLogos-32 contains photos depicting logos and is meant for the evaluation of multi-class logo detection/recognition as well as logo retrieval methods on real-world images. It consists of 8240 images downloaded from Flickr.

24. TME Motorway Dataset

    30000+ frames with vehicle rear annotation and classification (car and trucks) on motorway/highway sequences. Annotation semi-automatically generated using laser-scanner data. Distance estimation and consistent target ID over time available.

25. PHOS (Color Image Database for illumination invariant feature selection)

    Phos is a color image database of 15 scenes captured under different illumination conditions. More particularly, every scene of the database contains 15 different images: 9 images captured under various strengths of uniform illumination, and 6 images under different degrees of non-uniform illumination. The images contain objects of different shape, color and texture and can be used for illumination invariant feature detection and selection.

26. CaliforniaND: An Annotated Dataset For Near-Duplicate Detection In Personal Photo Collections

    California-ND contains 701 photos taken directly from a real user's personal photo collection, including many challenging non-identical near-duplicate cases, without the use of artificial image transformations. The dataset is annotated by 10 different subjects, including the photographer, regarding near duplicates.

27. USPTO Algorithm Challenge, Detecting Figures and Part Labels in Patents

    Contains drawing pages from US patents with manually labeled figure and part labels.

28. Abnormal Objects Dataset

    Contains 6 object categories similar to object categories in Pascal VOC that are suitable for studying the abnormalities stemming from objects.

29. Human detection and tracking using RGB-D camera

    Collected in a clothing store. Captured with Kinect (640*480, about 30fps)*

30. Multi-Task Facial Landmark (MTFL) dataset

    This dataset contains 12,995 face images collected from the Internet. The images are annotated with (1) five facial landmarks, (2) attributes of gender, smiling, wearing glasses, and head pose.

31. WIDER FACE: A Face Detection Benchmark

    WIDER FACE dataset is a face detection benchmark dataset with images selected from the publicly available WIDER dataset. It contains 32,203 images and 393,703 face annotations.

32. PIROPO Database: People in Indoor ROoms with Perspective and Omnidirectional cameras

    Multiple sequences recorded in two different indoor rooms, using both omnidirectional and perspective cameras, containing people in a variety of situations (people walking, standing, and sitting). Both annotated and non-annotated sequences are provided, where ground truth is point-based. In total, more than 100,000 annotated frames are available.

Classification

1. PASCAL VOC 2009 dataset

   Classification/Detection Competitions, Segmentation Competition, Person Layout Taster Competition datasets

2. Caltech

   Cars, Motorcycles, Airplanes, Faces, Leaves, Backgrounds

3. Caltech 101

Pictures of objects belonging to 101 categories

4. Caltech 256

   Pictures of objects belonging to 256 categories

5. ETHZ Shape Classes

   A dataset for testing object class detection algorithms. It contains 255 test images and features five diverse shape-based classes (apple logos, bottles, giraffes, mugs, and swans).

6. Flower classification data sets

   17 Flower Category Dataset

7. Animals with attributes

   A dataset for Attribute Based Classification. It consists of 30475 images of 50 animals classes with six pre-extracted feature representations for each image.

8. Stanford Dogs Dataset

   Dataset of 20,580 images of 120 dog breeds with bounding-box annotation, for fine-grained image categorization.

9. Video classification USAA dataset

   The USAA dataset includes 8 different semantic class videos which are home videos of social occassions which feature activities of group of people. It contains around 100 videos for training and testing respectively. Each video is labeled by 69 attributes. The 69 attributes can be broken down into five broad classes: actions, objects, scenes, sounds, and camera movement.

10. McGill Real-World Face Video Database

    This database contains 18000 video frames of 640x480 resolution from 60 video sequences, each of which recorded from a different subject (31 female and 29 male).

11. e-Lab Video Data Set

    Video data sets to train machines to recognise objects in our environment. e-VDS35 has 35 classes and a total of 2050 videos of roughly 10 seconds each.

Recognition

1. Face and Gesture Recognition Working Group FGnet

   Face and Gesture Recognition Working Group FGnet

2. Feret

   Face and Gesture Recognition Working Group FGnet

3. PUT face

   9971 images of 100 people

4. Labeled Faces in the Wild

   A database of face photographs designed for studying the problem of unconstrained face recognition

5. Urban scene recognition

   Traffic Lights Recognition, Lara's public benchmarks.

6. PubFig: Public Figures Face Database

   The PubFig database is a large, real-world face dataset consisting of 58,797 images of 200 people collected from the internet. Unlike most other existing face datasets, these images are taken in completely uncontrolled situations with non-cooperative subjects.

7. YouTube Faces

   The data set contains 3,425 videos of 1,595 different people. The shortest clip duration is 48 frames, the longest clip is 6,070 frames, and the average length of a video clip is 181.3 frames.

8. MSRC-12: Kinect gesture data set

   The Microsoft Research Cambridge-12 Kinect gesture data set consists of sequences of human movements, represented as body-part locations, and the associated gesture to be recognized by the system.

9. QMUL underGround Re-IDentification (GRID) Dataset

   This dataset contains 250 pedestrian image pairs + 775 additional images captured in a busy underground station for the research on person re-identification.

10. Person identification in TV series

Face tracks, features and shot boundaries from our latest CVPR 2013 paper. It is obtained from 6 episodes of Buffy the Vampire Slayer and 6 episodes of Big Bang Theory.

11. ChokePoint Dataset

    ChokePoint is a video dataset designed for experiments in person identification/verification under real-world surveillance conditions. The dataset consists of 25 subjects (19 male and 6 female) in portal 1 and 29 subjects (23 male and 6 female) in portal 2.

12. Hieroglyph Dataset

    Ancient Egyptian Hieroglyph Dataset.

13. Rijksmuseum Challenge Dataset: Visual Recognition for Art Dataset

    Over 110,000 photographic reproductions of the artworks exhibited in the Rijksmuseum (Amsterdam, the Netherlands). Offers four automatic visual recognition challenges consisting of predicting the artist, type, material and creation year. Includes a set of baseline features, and offer a baseline based on state-of-the-art image features encoded with the Fisher vector.

14. The OU-ISIR Gait Database, Treadmill Dataset

    Treadmill gait datasets composed of 34 subjects with 9 speed variations, 68 subjects with 68 subjects, and 185 subjects with various degrees of gait fluctuations.

15. The OU-ISIR Gait Database, Large Population Dataset

    Large population gait datasets composed of 4,016 subjects.

16. Pedestrian Attribute Recognition At Far Distance

    Large-scale PEdesTrian Attribute (PETA) dataset, covering more than 60 attributes (e.g. gender, age range, hair style, casual/formal) on 19000 images.

17. FaceScrub Face Dataset

    The FaceScrub dataset is a real-world face dataset comprising 107,818 face images of 530 male and female celebrities detected in images retrieved from the Internet. The images are taken under real-world situations (uncontrolled conditions). Name and gender annotations of the faces are included.

18. Depth-Based Person Identification

Depth-Based Person Identification from Top View Dataset.

Tracking

- Dataset-AMP: Luka Čehovin Zajc; Alan Lukežič; Aleš Leonardis; Matej Kristan. "Beyond Standard Benchmarks: Parameterizing Performance Evaluation in Visual Object Tracking." ICCV (2017).
  [paper(http://openaccess.thecvf.com/content*ICCV*2017/papers/Zajc*Beyond*Standard*Benchmarks*ICCV*2017*paper.pdf)]
- Dataset-Nfs: Hamed Kiani Galoogahi, Ashton Fagg, Chen Huang, Deva Ramanan and Simon Lucey.
  "Need for Speed: A Benchmark for Higher Frame Rate Object Tracking." ICCV (2017)
  [paper(http://openaccess.thecvf.com/content*ICCV*2017/papers/Galoogahi*Need*for*Speed*ICCV*2017*paper.pdf)]
  [supp(http://openaccess.thecvf.com/content*ICCV*2017/supplemental/Galoogahi-*Need*for*Speed*ICCV*2017*supplemental.pdf)]
  [project(http://ci2cv.net/nfs/index.html)]
- Dataset-DTB70: Siyi Li, Dit-Yan Yeung.
  "Visual Object Tracking for Unmanned Aerial Vehicles: A Benchmark and New Motion Models." AAAI (2017)
  [paper(http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14338/14292)]
  [project(https://github.com/flyers/drone-tracking)]
  [dataset(https://www.dropbox.com/s/s1fj99s2six4lrs/DTB70.tar.gz?dl=0)]
- Dataset-UAV123: Matthias Mueller, Neil Smith and Bernard Ghanem.
  "A Benchmark and Simulator for UAV Tracking." ECCV (2016)
  [paper(https://ivul.kaust.edu.sa/Documents/Publications/2016/A%20Benchmark%20and%20Simulator%20for%20UAV%20Tracking.pdf)]
  [project(https://ivul.kaust.edu.sa/Pages/pub-benchmark-simulator-uav.aspx)]

[dataset(https://ivul.kaust.edu.sa/Pages/Dataset-UAV123.aspx)]

- Dataset-TColor-128: Pengpeng Liang, Erik Blasch, Haibin Ling.
  "Encoding color information for visual tracking: Algorithms and benchmark." TIP (2015)
  [paper(http://www.dabi.temple.edu/ [project(http://www.dabi.temple.edu/ [dataset(http://www.dabi.temple.edu/Dataset-NUS-PRO: Annan Li, Min Lin, Yi Wu, Ming-Hsuan Yang, and Shuicheng Yan.
- "NUS-PRO: A New Visual Tracking Challenge." PAMI (2015)
  [paper(http://faculty.ucmerced.edu/mhyang/papers/pami15*nus*pro.pdf)]
  [project(https://sites.google.com/site/li00annan/nus-pro)]
  [Data*360*(https://d9fca6.lc.yunpan.cn/lk/cqKIc6DU3t2eJ)(code:bf28)]
  [Data*baidu*](https://pan.baidu.com/s/1pJHvbSn#list/path=%2F)]
  [View*360*(https://6aa275.lc.yunpan.cn/lk/cqK479PfzDrPX)(code:515a)]
  [View*baidu*](https://pan.baidu.com/s/1hqKXcuK)]

2. Dataset-PTB: Shuran Song and Jianxiong Xiao.
   "Tracking Revisited using RGBD Camera: Unified Benchmark and Baselines." ICCV (2013)
   [paper(http://vision.princeton.edu/projects/2013/tracking/paper.pdf)]
   [project(http://tracking.cs.princeton.edu/)]
   [5 validation(http://tracking.cs.princeton.edu/ValidationSet.zip)]
   [95 evaluation(http://tracking.cs.princeton.edu/EvaluationSet.tgz)]
3. Dataset-ALOV300+: Arnold W. M. Smeulders, Dung M. Chu, Rita Cucchiara, Simone Calderara, Afshin Dehghan, Mubarak Shah.
   "Visual Tracking: An Experimental Survey." PAMI (2014)
   [paper(http://crcv.ucf.edu/papers/Tracking*Survey.pdf)]*
   [project(http://imagelab.ing.unimore.it/dsm/)]
   Mirror Link:ALOV300(http://crcv.ucf.edu/people/phd*students/afshin/ALOV300/Frames.zip)*

Mirror Link:ALOV300(http://crcv.ucf.edu/people/phd*students/af-shin/ALOV300/GT.zip)*

4. OTB2013: Wu, Yi, Jongwoo Lim, and Minghsuan Yang. "Online Object Tracking: A Benchmark." CVPR (2013). [paper(http://faculty.ucmerced.edu/mhyang/papers/cvpr13*benchmark.pdf)]*
   - OTB2015: Wu, Yi, Jongwoo Lim, and Minghsuan Yang.
5. "Object Tracking Benchmark." TPAMI (2015). [paper(http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7001050&tag=1)] [project(http://cvlab.hanyang.ac.kr/tracker*benchmark/index.html)]*
   - Dataset-VOT:
6. [project(http://www.votchallenge.net/)]
7. VOT13paperICCV(http://www.votchallenge.net/vot2013/Download/vot2013paper.pdf)The Visual Object Tracking VOT2013 challenge results
8. [VOT14paperECCV(http://www.votchallenge.net/vot2014/download/vot2014paper.pdf)]The Visual Object Tracking VOT2014 challenge results
9. [VOT15paperICCV(http://www.votchallenge.net/vot2015/download/vot2015paper.pdf)]The Visual Object Tracking VOT2015 challenge results
10. [VOT16paperECCV(http://www.votchallenge.net/vot2016/download/vot2016paper.pdf)]The Visual Object Tracking VOT2016 challenge results
11. [VOT17paperECCV(http://openaccess.thecvf.com/contentICCV2017workshops/papers/w28/KristanTheVisualObjectICCV2017paper.pdf)]The Visual Object Tracking VOT2017 challenge results

Segmentation

1. Image Segmentation with A Bounding Box Prior dataset
   Ground truth database of 50 images with: Data, Segmentation, Labelling - Lasso, Labelling - Rectangle
2. PASCAL VOC 2009 dataset

Classification/Detection Competitions, Segmentation Competition, Person Layout Taster Competition datasets

3. Motion Segmentation and OBJCUT data

   Cows for object segmentation, Five video sequences for motion segmentation

4. Geometric Context Dataset

   Geometric Context Dataset: pixel labels for seven geometric classes for 300 images

5. Crowd Segmentation Dataset

   This dataset contains videos of crowds and other high density moving objects. The videos are collected mainly from the BBC Motion Gallery and Getty Images website. The videos are shared only for the research purposes. Please consult the terms and conditions of use of these videos from the respective websites.

6. CMU-Cornell iCoseg Dataset

   Contains hand-labelled pixel annotations for 38 groups of images, each group containing a common foreground. Approximately 17 images per group, 643 images total.

7. Segmentation evaluation database

   200 gray level images along with ground truth segmentations

8. The Berkeley Segmentation Dataset and Benchmark

   Image segmentation and boundary detection. Grayscale and color segmentations for 300 images, the images are divided into a training set of 200 images, and a test set of 100 images.

9. Weizmann horses

   328 side-view color images of horses that were manually segmented. The images were randomly collected from the WWW.

10. Saliency-based video segmentation with sequentially updated priors

    10 videos as inputs, and segmented image sequences as ground-truth

11. Daimler Urban Segmentation Dataset

    The dataset consists of video sequences recorded in urban traffic. The dataset consists of 5000 rectified stereo image pairs. 500 frames come with pixel-level semantic class

annotations into 5 classes: ground, building, vehicle, pedestrian, sky. Dense disparity maps are provided as a reference.

12. DAVIS: Densely Annotated VIdeo Segmentation

    A Benchmark Dataset and Evaluation Methodology for Video Object Segmentation.

Foreground/Background

1. Wallflower Dataset

   For evaluating background modelling algorithms

2. Foreground/Background Microsoft Cambridge Dataset

   Foreground/Background segmentation and Stereo dataset from Microsoft Cambridge

3. Stuttgart Artificial Background Subtraction Dataset

   The SABS (Stuttgart Artificial Background Subtraction) dataset is an artificial dataset for pixel-wise evaluation of background models.

4. Image Alpha Matting Dataset

   Image Alpha Matting Dataset.

5. LASIESTA: Labeled and Annotated Sequences for Integral Evaluation of SegmenTation Algorithms

   LASIESTA is composed by many real indoor and outdoor sequences organized in diferent categories, each of one covering a specific challenge in moving object detection strategies.

Saliency Detection (source)

1. AIM

   120 Images / 20 Observers (Neil D. B. Bruce and John K. Tsotsos 2005).

2. LeMeur

   27 Images / 40 Observers (O. Le Meur, P. Le Callet, D. Barba and D. Thoreau 2006).

3. Kootstra

100 Images / 31 Observers (Kootstra, G., Nederveen, A. and de Boer, B. 2008).

4. DOVES

   101 Images / 29 Observers (van der Linde, I., Rajashekar, U., Bovik, A.C., Cormack, L.K. 2009).

5. Ehinger

   912 Images / 14 Observers (Krista A. Ehinger, Barbara Hidalgo-Sotelo, Antonio Torralba and Aude Oliva 2009).

6. NUSEF

   758 Images / 75 Observers (R. Subramanian, H. Katti, N. Sebe1, M. Kankanhalli and T-S. Chua 2010).

7. JianLi

   235 Images / 19 Observers (Jian Li, Martin D. Levine, Xiangjing An and Hangen He 2011).

8. Extended Complex Scene Saliency Dataset (ECSSD)

   ECSSD contains 1000 natural images with complex foreground or background. For each image, the ground truth mask of salient object(s) is provided.

Video Surveillance

1. CAVIAR

   For the CAVIAR project a number of video clips were recorded acting out the different scenarios of interest. These include people walking alone, meeting with others, window shopping, entering and exitting shops, fighting and passing out and last, but not least, leaving a package in a public place.

2. ViSOR

   ViSOR contains a large set of multimedia data and the corresponding annotations.

3. CUHK Crowd Dataset

   474 video clips from 215 crowded scenes, with ground truth on group detection and video classes.?

4. TImes Square Intersection (TISI) Dataset

   A busy outdoor dataset for research on visual surveillance.

5. Educational Resource Centre (ERCe) Dataset

   An indoor dataset collected from a university campus for physical event understanding of long video streams.

6. PIROPO Database: People in Indoor ROoms with Perspective and Omnidirectional cameras

   Multiple sequences recorded in two different indoor rooms, using both omnidirectional and perspective cameras, containing people in a variety of situations (people walking, standing, and sitting). Both annotated and non-annotated sequences are provided, where ground truth is point-based. In total, more than 100,000 annotated frames are available.

Multiview

1. 3D Photography Dataset

   Multiview stereo data sets: a set of images

2. Multi-view Visual Geometry group's data set

   Dinosaur, Model House, Corridor, Aerial views, Valbonne Church, Raglan Castle, Kapel sequence

3. Oxford reconstruction data set (building reconstruction)

   Oxford colleges

4. Multi-View Stereo dataset (Vision Middlebury)

   Temple, Dino

5. Multi-View Stereo for Community Photo Collections

   Venus de Milo, Duomo in Pisa, Notre Dame de Paris

6. IS-3D Data

   Dataset provided by Center for Machine Perception

7. CVLab dataset

CVLab dense multi-view stereo image database

8. 3D Objects on Turntable

   Objects viewed from 144 calibrated viewpoints under 3 different lighting conditions

9. Object Recognition in Probabilistic 3D Scenes

   Images from 19 sites collected from a helicopter flying around Providence, RI. USA. The imagery contains approximately a full circle around each site.

10. Multiple cameras fall dataset

    24 scenarios recorded with 8 IP video cameras. The first 22 first scenarios contain a fall and confounding events, the last 2 ones contain only confounding events.

11. CMP Extreme View Dataset

    15 wide baseline stereo image pairs with large viewpoint change, provided ground truth homographies.

12. KTH Multiview Football Dataset II

    This dataset consists of 8000+ images of professional footballers during a match of the Allsvenskan league. It consists of two parts: one with ground truth pose in 2D and one with ground truth pose in both 2D and 3D.

13. Disney Research light field datasets

    This dataset includes: camera calibration information, raw input images we have captured, radially undistorted, rectified, and cropped images, depth maps resulting from our reconstruction and propagation algorithm, depth maps computed at each available view by the reconstruction algorithm without the propagation applied.

14. CMU Panoptic Studio Dataset

    Multiple people social interaction dataset captured by 500+ synchronized video cameras, with 3D full body skeletons and calibration data.

15. 4D Light Field Dataset

    24 synthetic scenes. Available data per scene: 9x9 input images (512x512x3) , ground truth (disparity and depth), camera parameters, disparity ranges, evaluation masks.

Action

1. UCF Sports Action Dataset

   This dataset consists of a set of actions collected from various sports which are typically featured on broadcast television channels such as the BBC and ESPN. The video sequences were obtained from a wide range of stock footage websites including BBC Motion gallery, and GettyImages.

2. UCF Aerial Action Dataset

   This dataset features video sequences that were obtained using a R/C-controlled blimp equipped with an HD camera mounted on a gimbal.The collection represents a diverse pool of actions featured at different heights and aerial viewpoints. Multiple instances of each action were recorded at different flying altitudes which ranged from 400-450 feet and were performed by different actors.

3. UCF YouTube Action Dataset

   It contains 11 action categories collected from YouTube.

4. Weizmann action recognition

   Walk, Run, Jump, Gallop sideways, Bend, One-hand wave, Two-hands wave, Jump in place, Jumping Jack, Skip.

5. UCF50

   UCF50 is an action recognition dataset with 50 action categories, consisting of realistic videos taken from YouTube.

6. ASLAN

   The Action Similarity Labeling (ASLAN) Challenge.

7. MSR Action Recognition Datasets

   The dataset was captured by a Kinect device. There are 12 dynamic American Sign Language (ASL) gestures, and 10 people. Each person performs each gesture 2-3 times.

8. KTH Recognition of human actions

Contains six types of human actions (walking, jogging, running, boxing, hand waving and hand clapping) performed several times by 25 subjects in four different scenarios: outdoors, outdoors with scale variation, outdoors with different clothes and indoors.

9. Hollywood-2 Human Actions and Scenes dataset

   Hollywood-2 datset contains 12 classes of human actions and 10 classes of scenes distributed over 3669 video clips and approximately 20.1 hours of video in total.

10. Collective Activity Dataset

    This dataset contains 5 different collective activities : crossing, walking, waiting, talking, and queueing and 44 short video sequences some of which were recorded by consumer hand-held digital camera with varying view point.

11. Olympic Sports Dataset

    The Olympic Sports Dataset contains YouTube videos of athletes practicing different sports.

12. SDHA 2010

    Surveillance-type videos

13. VIRAT Video Dataset

    The dataset is designed to be realistic, natural and challenging for video surveillance domains in terms of its resolution, background clutter, diversity in scenes, and human activity/event categories than existing action recognition datasets.

14. HMDB: A Large Video Database for Human Motion Recognition

    Collected from various sources, mostly from movies, and a small proportion from public databases, YouTube and Google videos. The dataset contains 6849 clips divided into 51 action categories, each containing a minimum of 101 clips.

15. Stanford 40 Actions Dataset

    Dataset of 9,532 images of humans performing 40 different actions, annotated with bounding-boxes.

16. 50Salads dataset

Fully annotated dataset of RGB-D video data and data from accelerometers attached to kitchen objects capturing 25 people preparing two mixed salads each (4.5h of annotated data). Annotated activities correspond to steps in the recipe and include phase (pre-/ core-/ post) and the ingredient acted upon.

17. Penn Sports Action

    The dataset contains 2326 video sequences of 15 different sport actions and human body joint annotations for all sequences.

18. CVRR-HANDS 3D

    A Kinect dataset for hand detection in naturalistic driving settings as well as a challenging 19 dynamic hand gesture recognition dataset for human machine interfaces.

19. TUM Kitchen Data Set

    Observations of several subjects setting a table in different ways. Contains videos, motion capture data, RFID tag readings,...

20. TUM Breakfast Actions Dataset

21. This dataset comprises of 10 actions related to breakfast preparation, performed by 52 different individuals in 18 different kitchens.

22. MPII Cooking Activities Dataset

    Cooking Activities dataset.

23. GTEA Gaze+ Dataset

    This dataset consists of seven meal-preparation activities, each performed by 10 subjects. Subjects perform the activities based on the given cooking recipes.

24. UTD-MHAD: multimodal human action recogniton dataset

    The dataset consists of four temporally synchronized data modalities. These modalities include RGB videos, depth videos, skeleton positions, and inertial signals (3-axis acceleration and 3-axis angular velocity) from a Kinect RGB-D camera and a wearable inertial sensor for a comprehensive set of 27 human actions.

Human pose/Expression

1. AFEW (Acted Facial Expressions In The Wild)/SFEW (Static Facial Expressions In The Wild)

   Dynamic temporal facial expressions data corpus consisting of close to real world environment extracted from movies.

2. Expression in-the-Wild (ExpW) Dataset

   Contains 91,793 faces manually labeled with expressions. Each of the face images was manually annotated as one of the seven basic expression categories: â€œangryâ€, â€œdisgustâ€, â€œfearâ€, â€œhappyâ€, â€œsadâ€, â€œsurpriseâ€, or â€œneutralâ€.

3. ETHZ CALVIN Dataset

   CALVIN research group datasets

4. HandNet (annotated depth images of articulating hands)

   This dataset includes 214971 annotated depth images of hands captured by a RealSense RGBD sensor of hand poses. Annotations: per pixel classes, 6D fingertip pose, heatmap. Images -> Train: 202198, Test: 10000, Validation: 2773. Recorded at GIP Lab, Technion.

5. 3D Human Pose Estimation

   Depth videos + ground truth human poses from 2 viewpoints to improve 3D human pose estimation.

Medical

1. VIP Laparoscopic / Endoscopic Dataset

   Collection of endoscopic and laparoscopic (mono/stereo) videos and images

2. Mouse Embryo Tracking Database

   DB Contains 100 examples with the uncompressed frames, up to the 10th frame after the appearance of the 8th cell; a text file with the trajectories of all the cells, from appearance to division; a movie file showing the trajectories of the cells.

3. FIRE Fundus Image Registration Dataset

   134 retinal image pairs and ground truth for registration.

Misc

1. Zurich Buildings Database

   ZuBuD Image Database contains over 1005 images about Zurich city building.
2. Color Name Data Sets
3. Mall dataset

   The mall dataset was collected from a publicly accessible webcam for crowd counting and activity profiling research.
4. QMUL Junction Dataset

   A busy traffic dataset for research on activity analysis and behaviour understanding.
5. Miracl-VC1

   Miracl-VC1 is a lip-reading dataset including both depth and color images. Fifteen speakers positioned in the frustum of a MS Kinect sensor and utter ten times a set of ten words and ten phrases.
6. NYU Symmetry Database

   The mirror symmetry database contains 176 single-symmetry and 63 multiple-symmetry images (.png files) with accompanying ground-truth annotations (.mat files).
7. RGB-W: When Vision Meets Wireless

   Data with the wireless signal emitted by individuals' cell phones, referred to as RGB-W.

Challenge

1. Microsoft COCO Image Captioning Challenge

   https://competitions.codalab.org/competitions/3221
2. ImageNet Large Scale Visual Recognition Challenge

http://www.image-net.org/

3. COCO 2017 Detection Challenge
   http://cocodataset.org/#detections-challenge2017
4. Visual Domain Adaptation (VisDA2017) Segmentation Challenge
   https://competitions.codalab.org/competitions/17054
5. The PASCAL Visual Object Classes Homepage
   http://host.robots.ox.ac.uk/pascal/VOC/
6. YouTube-8M Large-Scale Video Understanding
   https://research.google.com/youtube8m/workshop.html
7. joint COCO and Places Challenge
   https://places-coco2017.github.io/
8. Places Challenge 2017: Deep Scene Understanding is held jointly with COCO Challenge at ICCV'17
   http://placeschallenge.csail.mit.edu/
9. COCO Challenges.
   http://cocodataset.org/#home
10. VQA Challenge 2017
    http://visualqa.org/
11. The Joint Video and Language Understanding Workshop: MovieQA and The Large Scale Movie Description Challenge (LSMDC), at ICCV 2017
    https://sites.google.com/site/describingmovies/challenge
12. Microsoft Multimedia Challenge (2017)
    http://ms-multimedia-challenge.com/2017/challenge
13. MOTChallenge: The Multiple Object Tracking Benchmark
    https://motchallenge.net/
14. Visual Domain Adaptation Challenge
    http://ai.bu.edu/visda-2017/
15. MegaFace and MF2: Million-Scale Face Recognition

http://megaface.cs.washington.edu/

16. Facial Keypoints Detection

    https://www.kaggle.com/c/facial-keypoints-detection

17. The VOT challenges Visual Object Tracking

    http://www.votchallenge.net/

18. Large-scale Scene Understanding Challenge. SCENE CLASSIFICATION, SEGMENTATION, SALIENCY PREDICTION

    http://lsun.cs.princeton.edu/2017/

19. AI Challenger·全球 AI 挑战赛 图像中文描述，人体骨骼关键点，场景分类

    https://challenger.ai/

20. 2016 上海 BOT 大数据应用大赛

    http://www.zhishu51.com/Activity/bot

创业公司

1. 旷视科技：让机器看懂世界

   https://megvii.com/

2. 云从科技：源自计算机视觉之父的人脸识别技术

   http://www.cloudwalk.cn/

3. 格林深瞳：让计算机看懂世界

   http://www.deepglint.com/

4. 北京陌上花科技有限公司：人工智能计算机视觉引擎

   http://www.dressplus.cn/

5. 依图科技：与您一起构建计算机视觉的未来

   http://www.yitutech.com/

6. 码隆科技：最时尚的人工智能

   https://www.malong.com/

7. Linkface 脸云科技：全球领先的人脸识别技术服务

https://www.linkface.cn/

8. 速感科技：让机器人认识世界，用机器人改变世界

   http://www.qfeeltech.com/

9. 图森： 中国自动驾驶商业化领跑者

   http://www.tusimple.com/

10. Sense TIme 商汤科技：教会计算机看懂这个世界

    https://www.sensetime.com/

11. 图普科技：专注于图像识别

    https://us.tuputech.com/?from=gz

12. 亮风台： 专注增强现实，引领人机交互

    https://www.hiscene.com/

13. 中科视拓 ： 知人识面辨万物，开源赋能共发展

    http://www.seetatech.com/

公众号

1. 视觉求索 thevisionseeker
2. 深度学习大讲堂 deeplearningclass
3. VALSE valse*wechat*

汇总不全面，欢迎补全和提建议，敬请关注 http://www.zhuanzhi.ai 和关注专知公众号，获取第一手 AI 相关知识

4. 

5.