* [ImageNet Dataset](http://www.image-net.org/)
* [Wikipedia, Computer Vision](https://en.wikipedia.org/wiki/Computer_vision)
* [Fei-Fei Li, J. Johnson, s. Yeung, Convolutional Neural Networks for Visual Recognition, Spring 2018](http://cs231n.stanford.edu/)
* [Convolutional Neural Networks (CNNs / ConvNets)](https://cs231n.github.io/convolutional-networks/)
* [J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, K. Murphy, Speed/accuracy trade-offs for modern convolutional object detectors, 2017](https://arxiv.org/abs/1611.10012)
* [R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, 2014](https://arxiv.org/abs/1311.2524)
* [R. Girshick, Fast R-CNN, 2015](https://arxiv.org/abs/1504.08083)
* [S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, 2015](https://arxiv.org/abs/1506.01497)
* [J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You Only Look Once: Unified, Real-Time Object Detection, 2015](https://arxiv.org/abs/1506.02640)
* [W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, A. C. Berg, SSD: Single Shot MultiBox Detector, 2015](https://arxiv.org/abs/1512.02325)
* [J. Long, E. Shelhamer, T. Darrell, Fully Convolutional Networks for Semantic Segmentation, 2015](https://arxiv.org/abs/1411.4038)
* [H. Noh, S. Hong, B. Han, Learning Deconvolution Network for Semantic Segmentation, 2015](https://arxiv.org/abs/1505.04366)
* [K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, 2017](https://arxiv.org/abs/1703.06870)
* [Show and Tell: A Neural Image Caption Generator](https://arxiv.org/pdf/1411.4555.pdf)
* [Show, Attend and Tell: Neural Image Caption Generation with Visual Attention](https://arxiv.org/pdf/1502.03044.pdf)
* [End-to-End Object Detection with Transformers](https://arxiv.org/pdf/2005.12872.pdf)
* [High Performance Visual Tracking with Siamese Region Proposal Network](https://openaccess.thecvf.com/content_cvpr_2018/papers/Li_High_Performance_Visual_CVPR_2018_paper.pdf)
* [Human Action Recognition and Prediction: A Survey](https://arxiv.org/pdf/1806.11230.pdf)
* [Two-Stream Convolutional Networks for Action Recognition in Videos](https://arxiv.org/abs/1406.2199)
* [Learning Spatiotemporal Features with 3D Convolutional Networks](https://arxiv.org/pdf/1412.0767.pdf)
* [Deep Learning on Image Denoising: An overview](https://arxiv.org/pdf/1912.13171.pdf)
* [Deep Learning for Image Super-resolution: A Survey](https://arxiv.org/pdf/1902.06068.pdf)
* [NIPS 2016 Tutorial: Generative Adversarial Networks](https://arxiv.org/pdf/1701.00160.pdf)
* CGAN: <https://arxiv.org/pdf/1411.1784.pdf>
* InfoGAN: <https://arxiv.org/pdf/1606.03657.pdf>
* DCGAN: <https://arxiv.org/pdf/1511.06434.pdf>
* WGAN: <https://arxiv.org/pdf/1701.07875.pdf>
* OpenAI Tutorial: <https://blog.openai.com/generative-models/>
* GAN Lab: A great visual source to learn GANs: <https://poloclub.github.io/ganlab/>
* GAN Zoo: <https://github.com/hindupuravinash/the-gan-zoo>