Reference List for Deep Learning

Garry

Last updated on February 20, 2017

Contents

1 Deep Learning Foundations

1.1 Books

1. Deep Learning(Goodfellow, Bengio, and Courville 2016) is a textbook on Deep Learning written by Ian Goodfellow, Yoshua Bengio, and Aaron Courville

1.2 Papers

1. Learning representations by back-propagating errors(Rumelhart, Hinton, and Williams 1988)
2. Batch normalization: Accelerating deep network training by reducing internal covariate shift(Ioffe and Szegedy 2015)
3. Highway networks(Srivastava, Greff, and Schmidhuber 2015)

2 Deep Learning for Images

2.1 Papers

1. Very Deep Convolutional Networks for Large-scale Image Recognition, **VGG**(Simonyan and Zisserman 2014)
2. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention(Xu et al. 2015)
3. Going deeper with convolutions(Szegedy et al. 2015)
4. Deep residual learning for image recognition(He et al. 2015)
5. Deep Visual-Semantic Alignments for Generating Image Descriptions(Karpathy and Fei-Fei 2015)

3 Deep Learning for Speech/Videos

3.1 Books

1. Automatic speech recognition: A deep learning approach(Yu and Deng 2014)

3.2 Papers

1. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups(Hinton et al. 2012)

4 Deep Learning for Texts

4.1 Papers

1. Generating Text with Recurrent Neural Networks(Sutskever, Martens, and Hinton 2011)
2. Distributed representations of words and phrases and their compositionality(Mikolov et al. 2013a)
3. Efficient Estimation of Word Representations in Vector Space(Mikolov et al. 2013b)
4. Linguistic Regularities in Continuous Space Word Representations(Mikolov, Yih, and Zweig 2013)

5 Online Resources

5.1 Courses

1. CS231n Convolutional Neural Networks for Visual Recognition, Stanford University
2. COS429 Fall 2014: Computer Vision, Princeton

5.2 Tutorial/Blog/Report/Slide

1. Restricted Boltzmann Machines (RBM), deeplearning.net
2. The Unreasonable Effectiveness of Recurrent Neural Networks, Andrej Karpathy blog
3. colah's blog
   (a) Understanding Convolutions
   (b) Visualizing MNIST: An Exploration of Dimensionality Reduction
   (c) Visual Information Theory
4. WILDML' BLOG
   (a) Implementing a Neural Network from Scratch in Python an Introduction
5. Learning Multiple Layers of Features from Tiny Images, Alex Krizhevsky

6 Open source code/tools

1. TensorFlow
   (a) Deep MNIST for Experts
2. cuda-convnet, cuda-convnet2
3. DyNet(Neubig et al. 2017)
4. Theano

7 Datasets

7.1 Computer Vision

1. Wild (LFW) dataset
2. The CIFAR-10 dataset, The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.

7.2 Speech/Videos

1. TIMIT Acoustic-Phonetic Continuous Speech Corpus

8 Math

8.1 Probability Theory

1. Probability theory: The logic of science(Jaynes 2003)

8.2 Linear Algebra

1. The Matrix Book,(Petersen, Pedersen, et al. 2008)
2. Linear Algebra. Revised English Ed. Trs. RA Silverman(SHILOV and SILVERMAN 1971)
3. Introduction to Linear Algebra, Fifth Edition (2016)

9 Programming

9.1 C/C++

1. The Definitive C++ Book Guide and List
2. CMake official documentation

9.2 Python

1. SciPy: Quickstart tutorial

10 Active Researchers

Deep Learning Research Groups

1. homepage, Alex Krizhevsky
2. Homepage, Hui Zhang, CMU

11 Research and Writing

Advice on Research and Writing:A collection of advice about how to do research and how to communicate effectively (primarily for computer scientists).

References

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*. http://www.deeplearningbook.org. MIT Press.

He, Kaiming et al. (2015). "Deep residual learning for image recognition". In: *arXiv preprint arXiv:1512.03385*.

Hinton, Geoffrey et al. (2012). "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups". In: *IEEE Signal Processing Magazine* 29.6, pp. 82–97.

Ioffe, Sergey and Christian Szegedy (2015). "Batch normalization: Accelerating deep network training by reducing internal covariate shift". In: *arXiv preprint arXiv:1502.03167*.

Jaynes, Edwin T (2003). *Probability theory: The logic of science*. Cambridge university press.

Karpathy, Andrej and Li Fei-Fei (2015). "Deep visual-semantic alignments for generating image descriptions". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3128–3137.

Mikolov, Tomas, Wen-tau Yih, and Geoffrey Zweig (2013). "Linguistic Regularities in Continuous Space Word Representations." In: *Hlt-naacl*. Vol. 13, pp. 746–751.

Mikolov, Tomas et al. (2013a). "Distributed representations of words and phrases and their compositionality". In: *Advances in neural information processing systems*, pp. 3111–3119.

Mikolov, Tomas et al. (2013b). "Efficient estimation of word representations in vector space". In: *arXiv preprint arXiv:1301.3781*.

Neubig, Graham et al. (2017). "DyNet: The Dynamic Neural Network Toolkit". In: *arXiv preprint arXiv:1701.03980*.

Petersen, Kaare Brandt, Michael Syskind Pedersen, et al. (2008). "The matrix cookbook". In: *Technical University of Denmark* 7, p. 15.

Rumelhart, David E, Geoffrey E Hinton, and Ronald J Williams (1988). "Learning representations by back-propagating errors". In: *Cognitive modeling* 5.3, p. 1.

SHILOV, Georgy Evgen'evich and Richard A SILVERMAN (1971). *Linear Algebra. Revised English Ed. Trs. RA Silverman*. Prentice-Hall.

Simonyan, Karen and Andrew Zisserman (2014). "Very deep convolutional networks for large-scale image recognition". In: *arXiv preprint arXiv:1409.1556*.

Srivastava, Rupesh Kumar, Klaus Greff, and Jürgen Schmidhuber (2015). "Highway networks". In: *arXiv preprint arXiv:1505.00387*.

Sutskever, Ilya, James Martens, and Geoffrey E Hinton (2011). "Generating text with recurrent neural networks". In: *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pp. 1017–1024.

Szegedy, Christian et al. (2015). "Going deeper with convolutions". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9.

Xu, Kelvin et al. (2015). "Show, attend and tell: Neural image caption generation with visual attention". In: *arXiv preprint arXiv:1502.03044* 2.3, p. 5.

Yu, Dong and Li Deng (2014). *Automatic speech recognition: A deep learning approach*. Springer.