

Deep Learning References

Pablo Mesejo
Inria Grenoble Rhône-Alpes
Perception team

April 4, 2017

Abstract

This document contains some potentially useful references to understand artificial neural networks (ANNs) and deep learning (DL) methods, at both theoretical and practical levels.

1 Textbooks and surveys about DL

- Schmidhuber, J. (2015). “Deep Learning in Neural Networks: An Overview”. *Neural Networks* 61: 85-117.
- Bengio, Y., LeCun, Y., and Hinton, G. (2015). “Deep Learning”. *Nature* 521: 436-44.

The authors of the previous review papers maintained a very interesting public controversy about giving credit to the pioneers of the field: <https://plus.google.com/100849856540000067209/posts/9BDtGwCDL7D>

- Goodfellow, I., Bengio, Y., and Courville, A. (2016). “Deep Learning”. <http://www.deeplearningbook.org/> and <https://github.com/HFTrader/DeepLearningBook>. The official webpage even offer lecture slides accompanying some chapters of the book.
- Bengio, Y., Courville, A., and Vincent, P. (2013). “Representation learning: A review and new perspectives”, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35 (8): 1798-1828.

- Arel, I., Rose, D.C., and Karnowski, T.P. (2010). “Deep Machine Learning - A New Frontier in Artificial Intelligence Research”. IEEE Computational Intelligence Magazine 5 (4): 13-18.
- Bengio, Y. (2009) “Learning deep architectures for AI”. Foundations and trends in Machine Learning 2 (1): 1-127

2 Introductory books and tutorials on ANNs

- Bishop, C.M. (1995) Neural Networks for Pattern Recognition, Oxford University Press.
- Haykin, S. (1999) Neural Networks: A Comprehensive Foundation, Prentice Hall.
- Bishop, C.M. (2006) Pattern Recognition and Machine Learning, Springer. Chapter 5 is dedicated to Neural Networks.
- “Neural Networks and Deep Learning” by Michael Nielsen: <http://neuralnetworksanddeeplearning.com/index.html>
- Tutorials on neural networks and deep learning by Quoc V. Le: <https://cs.stanford.edu/~quocle/tutorial1.pdf>, <https://cs.stanford.edu/~quocle/tutorial2.pdf>, and <http://www.trivedigaurav.com/blog/quoc-les-lectures-on-deep-learning/>

3 Some recommended references in specific subjects

3.1 Convolutional Neural Networks

- “Visualizing and Understanding Convolutional Networks” by Matthew D. Zeiler and Rob Fergus (2014)
- “Convolutional Neural Networks for Visual Recognition” (Stanford course given by Fei-Fei Li, Andrej Karpathy, and Justin Johnson, 2016): <http://cs231n.github.io/>

- “A beginner’s guide to understanding Convolutional Neural Networks” by Adit Deshpand <https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/>
- “Understanding Deep Convolutional Networks” by Stéphane Mallat (2016)
- “Convolutional Neural Networks” by Nando de Freitas (2015): https://www.youtube.com/watch?v=bEUX_56Lojc

3.2 Unsupervised Deep Learning

- “Generative Adversarial Networks” (2014) by Ian J. Goodfellow et al.
- “Auto-Encoding Variational Bayes” (2013) by Diederik P. Kingma and Max Welling.
- “Tutorial on Variational Autoencoders” (2016) by Carl Doersch.
- “NIPS 2016 Workshop on Adversarial Training”: <https://www.youtube.com/playlist?list=PLJscN9YDD1buxCitmej1pjJkR5PMhenTF>

3.3 Recurrent Neural Networks

- “Supervised Sequence Labelling with Recurrent Neural Networks” (2012) by Alex Graves.
- “A Critical Review of Recurrent Neural Networks for Sequence Learning” (2015) by Z.C. Lipton et al.
- Deep Natural Language Processing course offered at the University of Oxford: <https://github.com/oxford-cs-deepnlp-2017/lectures>
- “The Unreasonable Effectiveness of Recurrent Neural Networks” by Andrej Karpathy: <https://karpathy.github.io/2015/05/21/rnn-effectiveness/>
- “Understanding LSTM Networks” by Christopher Olah: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- “LSTM: A search space odyssey” (2016) by K. Greff et al.
- “Training Recurrent Neural Networks” (2012) by Ilya Sutskever

3.4 Reinforcement Learning

- “Reinforcement Learning: An Introduction” by Richard S. Sutton and Andrew G. Barto: <https://webdocs.cs.ualberta.ca/~sutton/book/the-book-2nd.html>
- David Silver’s course: <http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching.html>
- “Deep Reinforcement Learning: Pong from Pixels” by Andrej Karpathy: <https://karpathy.github.io/2016/05/31/r1/>
- Talks on Deep Reinforcement Learning by John Schulman: https://www.youtube.com/watch?v=aUrX-rP_ss4, and his Deep Reinforcement Learning course <http://rll.berkeley.edu/deeprlcourse/>.
- Andrew Ng’s Thesis: <http://rll.berkeley.edu/deeprlcourse/docs/ng-thesis.pdf>

4 More resources online

- Reading lists, survey papers, and most cited deep learning papers:
 - <http://deeplearning.net/reading-list/>
 - <https://github.com/terryum/awesome-deep-learning-papers>
 - <https://github.com/IshmaelBelghazi/Deep-Learning-Papers-Reading-Roadmap>
- Inria deep learning reading group sessions: <https://project.inria.fr/deeplearning/sessions/>
- Nando de Freitas’ talks: <https://www.youtube.com/user/ProfNandoDF/videos>
- Christopher Olah’s blog: <https://colah.github.io/>
- Andrej Karpathy’s blog: <https://karpathy.github.io/>
- Andrej Karpathy’s talks: https://www.youtube.com/channel/UCPk8m_r6fkUSYmvgCBwq-sw/videos

- Hugo Larochelle’s talks: <https://www.youtube.com/playlist?list=PL6Xpj9I5qXYEc0hn7TqghAJ6NAPrNmUBH>
- Adit Deshpande’s blog: <https://adeshpande3.github.io/>
- “Deep Learning” by Geoff Hinton (2015): <https://www.youtube.com/watch?v=IcOMKXAw5VA>
- “Introduction to neural nets and backpropagation” by Patrick Winston (2010): <https://www.youtube.com/watch?v=q0pm3BrIUfo>
- Deep Learning Summer School (Montreal, 2015): http://videlectures.net/deeplearning2015_montreal/
- Deep Learning Summer School (Montreal, 2016): http://videlectures.net/deeplearning2016_montreal/
- International Conference on Learning Representations (ICLR) 2016: http://videlectures.net/iclr2016_san_juan/
- International Conference on Machine Learning (ICML) 2016 Tutorials: <http://techtalks.tv/icml/2016/tutorials/>
- Neural Information Processing Systems (NIPS) 2016 Tutorials: <https://nips.cc/Conferences/2016/Schedule?type=Tutorial>
- “Scaling Up Deep Learning” by Yoshua Bengio (2014): http://videlectures.net/kdd2014_bengio_deep_learning/
- “Deep Learning” (slides by Geoff Hinton, Yoshua Bengio and Yann LeCun, NIPS’2015 tutorial) <http://www.iro.umontreal.ca/~bengioy/talks/DL-Tutorial-NIPS2015.pdf>
- “What’s Wrong with Deep Learning” (slides by Yann LeCun, CVPR’2015 keynote) <https://drive.google.com/file/d/0BxKBnD5y2M8NVHRiVXBn0VpiYUk>
- “Deep Learning Tutorial” (slides by Yann LeCun, ICML’2013 tutorial) <http://www.cs.nyu.edu/~yann/talks/lecun-ranzato-icml2013.pdf>
- Deep learning Udacity course: <https://classroom.udacity.com/courses/ud730/lessons/6370362152/concepts/63798118150923>

- Geoff Hinton’s course on Neural Networks for Machine Learning at Coursera: <https://www.coursera.org/learn/neural-networks>
- Andrew Ng’s course on Machine Learning at Coursera: <https://www.coursera.org/learn/machine-learning>
- “Backpropagation tutorial” by Manfred Zabarauskas (2011): <http://blog.zabarauskas.com/backpropagation-tutorial/>
- Introduction to deep neural networks: <http://deeplearning4j.org/neuralnet-overview.html>
- Neural Networks terminology: <http://www.asimovinstitute.org/neural-network-zoo/>
- A Guide to Deep Learning: <http://yerevann.com/a-guide-to-deep-learning/>
- Deep Learning course: lecture slides and lab notebooks. This course is being taught at as part of Master Datascience Paris Saclay: <https://m2dsupsdclass.github.io/lectures-labs/>

5 Some important papers...

- “A learning algorithm for Boltzmann machines” (1985), D.H. Ackley et al.
- “Learning representations by back-propagating errors” (1986), D.E. Rumelhart et al.
- “Learning internal representations by error-propagation” (1986), D.E. Rumelhart et al.
- “Backpropagation applied to handwritten zip code recognition” (1989), Y. LeCun et al.
- “Learning long-term dependencies with gradient descent is difficult” (1994), Y. Bengio et al.
- “Long short-term memory” (1997), S. Hochreiter and J. Schmidhuber

- “Gradient-based learning applied to document recognition” (1998), Y. LeCun et al.
- “Evolving Artificial Neural Networks” (1999), X. Yao
- “Learning to forget: Continual prediction with LSTM” (2000), F.A. Gers et al.
- “A fast learning algorithm for deep belief nets” (2006), G.E. Hinton et al.
- “Reducing the dimensionality of data with neural networks” (2006), G.E. Hinton and R.R. Salakhutdinov
- “To recognize shapes, first learn to generate images” (2007), G.E. Hinton
- “Learning Multiple Layers of Representation” (2007), G.E. Hinton
- “Greedy layer-wise training of deep networks” (2007), Y. Bengio et al.
- “What is the best multi-stage architecture for object recognition?” (2009), K. Jarrett et al.
- “A novel connectionist system for unconstrained handwriting recognition” (2009), A. Graves et al.
- “Rectified linear units improve restricted boltzmann machines” (2010), V. Nair and G.E. Hinton
- “Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion” (2010), P. Vincent et al.
- “Why does unsupervised pre-training help deep learning” (2010), D. Erhan et al.
- “Understanding the difficulty of training deep feedforward neural networks” (2010), X. Glorot and Y. Bengio
- “Deep sparse rectifier neural networks” (2011), X. Glorot et al.
- “Improving neural networks by preventing co-adaptation of feature detectors” (2012), G.E. Hinton et al.

- “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups” (2012), G.E. Hinton et al.
- “Efficient backprop” (2012), Y. LeCun et al.
- “Multi-column deep neural networks for image classification” (2012), D. Ciregan et al.
- “ImageNet classification with deep convolutional neural networks” (2012), A. Krizhevsky et al.
- “Large scale distributed deep networks” (2012), J. Dean et al.
- “Maxout networks” (2013), I. Goodfellow et al.
- “Network in network” (2013), M. Lin et al.
- “How transferable are features in deep neural networks?” (2014), J. Yosinski et al.
- “Dropout: A simple way to prevent neural networks from overfitting” (2014), N. Srivastava et al.
- “Where do features come from?” (2014), G.E. Hinton
- “Very deep convolutional networks for large-scale image recognition” (2014), K. Simonyan and A. Zisserman
- “OverFeat: Integrated recognition, localization and detection using convolutional networks” (2014), P. Sermanet et al.
- “Rich feature hierarchies for accurate object detection and semantic segmentation” (2014), R. Girshick et al.
- “Going deeper with convolutions” (2015), C. Szegedy et al.
- “Deep neural networks are easily fooled: High confidence predictions for unrecognizable images” (2015), A. Nguyen et al.
- “Fast R-CNN” (2015), R. Girshick
- “Fully convolutional networks for semantic segmentation” (2015), J. Long et al.

- “Deep Visual-Semantic Alignments for Generating Image Descriptions” (2015), A. Karpathy and L. Fei-Fei
- “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift” (2015), S. Ioffe and C. Szegedy
- “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks” (2016), S. Ren et al.
- “Deep residual learning for image recognition” (2016), K. He et al.
- “Spatial Transformer Networks” (2016), M. Jaderberg et al.
- “Region-based convolutional networks for accurate object detection and segmentation” (2016), R. Girshick et al.
- “Understanding deep learning requires re-thinking generalization” (2016), C. Zhang et al.

6 Libraries and simulators

- Keras: <https://keras.io/>
- TensorFlow: <https://www.tensorflow.org/>
- Theano: <http://deeplearning.net/software/theano/>
- Torch: <http://torch.ch/>
- Caffe: <http://caffe.berkeleyvision.org/>
- Exercises in python: <https://github.com/syhw/DL4H>
- Animated plug-in to gain intuitions about how ANNs behave <http://playground.tensorflow.org/>
- Software links to many toolboxes: http://deeplearning.net/software_links/