บรรณานุกรม

BCB14 D. Bahdanau, K. Cho, and Y. Bengio. Neural machine translation by jointly learning to align and translate. 2014.

BLB17 A. Botev, G. Lever, and D. Barber. Nesterov's accelerated gradient and momentum as approximations to regularised update descent. In *2017 International Joint Conference on Neural Networks (IJCNN)*, 1899–1903. IEEE, 2017.

BV18 S. Boyd and L. Vandenberghe. *Introduction to applied linear algebra: vectors, matrices and least squares*. Cambridge University Press, 2018.

CVMBB14 K. Cho, B. Van M., D. Bahdanau, and Y. Bengio. On the properties of neural machine translation: encoder-decoder approaches. *Association for Computational Linguistics*, pages 103–111, 2014.

CGCB14 J. Chung, C. Gulcehre, K. Cho, and Y. Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014. Presented in NIPS 2014 Deep Learning and Representation Learning Workshop.

CV95 C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20:273–297, 1995.

HZRS16 K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

HZSS15 K. He, X. Zhang, R. Shaoqing, and J. Sun. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. *IEEE International Conference on Computer Vision*, pages 1026–1034, 2015.

Heb49 D. Hebb. The first stage of perception: growth of the assembly. *The Organization of Behavior*, pages 60–78, 1949.

HS97 S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

Hop82 J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. In *Proceeding National Academy of Sciences*, volume 79, 2554–2558. 1982.

IS15 S. Ioffe and C. Szegedy. Batch normalization: accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on Machine Learning*, 37:448–456, 2015.

KB14 D. Kingma and J. Ba. Adam: a method for stochastic optimization. 2014. Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego,. URL: http://arxiv.org/abs/1412.6980.

KSH12 A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25(2):1097–1105, 2012.

LBD+89 Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4):541–551, 1989.

LBBH98 Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

MP43 W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, pages 115–133, 1943.

MP69 M. Minsky and S. Papert. *Perceptrons: an introduction to computational geometry*. MIT Press, 1969.

Mor19 L. Moroney. Horses or humans dataset. 2019. URL: http://laurencemoroney.com/horses-or-humans-dataset.

Mor20 L. Moroney. Deeplearning.ai tensorflow developer professional certificate. Coursera, 2020. URL: https://www.coursera.org/professional-certificates/tensorflow-in-practice.

Ng20 A. Ng. Deep learning specialization. Coursera, 2020. URL: https://www.coursera.org/specializations/deep-learning.

PEZ+14 S. Pierre, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun. Overfeat: integrated recognition, localization and detection using convolutional networks. In *2nd International Conference on Learning Representations*. 2014.

RDGF16 J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: unified, real-time object detection. *IEEE Conference on Computer VIsion and Pattern Recognition*, pages 779–788, 2016.

Ros58 F. Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408, 1958.

SKP15 F. Schroff, D. Kalenichenko, and J. Philbin. Facenet: a unified embedding for face recognition and clustering. In *Conference on Computer Vision and Pattern Recognition*, 815–823. 2015.

SZ14 K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

SLJ+15 C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. 2015.

TYRW14 Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. Deepface: closing the gap to human-level performance in face verification. *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1701–1708, 2014.

TH12 T. Tieleman and G. Hinton. Lecture 6.5 - rmsprop: divide the gradient by a running average of its recent magnitude. COURSERA : Neural Networks for Machine Learning, 2012. URL: https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf.

Web20 K. Webster. Tensorflow 2 for deep learning specialization. Coursera : Imperial College London, 2020. URL: https://www.coursera.org/specializations/tensorflow2-deeplearning.

XB10 G. Xavier and Y. Bengio. Understaning the difficulty of training dep feedforward neural networks. *Aistats*, 9:249–256, 2010.

XBK+15 K. Xu, J. Ba, R. Kiros, K. Cho, A. C. Courville, R. Salakhutdinov, R. S. Zemel, and Y. Bengio. Show, attend and tell: neural image caption generation with visual attention. *CoRR*, 2015.