DAFTAR PUSTAKA

Alpaydin, E. (2010). Introduction to Machine Learning Second Edition. In *Introduction to Machine Learning*. https://doi.org/10.1007/978-1-62703-748-8_7

Chilimbi, T., Suzue, Y., Apacible, J., & Kalyanaraman, K. (2014). Project Adam: Building an Efficient and Scalable Deep Learning Training System. *Proceedings of the 11th USENIX Symposium on Operating Systems Design and Implementation*, 571–582. Retrieved from https://www.usenix.org/conference/osdi14/technical-sessions/presentation/chilimbi

Clevert, D.-A., Unterthiner, T., & Hochreiter, S. (2015). *Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)*. 1–14. Retrieved from http://arxiv.org/abs/1511.07289

Dean, J., & Corrado, G. S. (2012). Large Scale Distributed Deep Networks Jeffrey. *Cambridge University Press*, 1–9. https://doi.org/10.1109/ICDAR.2011.95

Ioffe, S., & Szegedy, C. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. Retrieved from http://arxiv.org/abs/1502.03167

Le, Q. V., Jiquan Ngiam, Adam Coates, Abhik Lahiri, Prochnow, B., & Andrew Y. Ng. (2016). On Optimization Methods for Deep Learning. *European Solid-State Circuits Conference*, *2016-Octob*, 129–132. https://doi.org/10.1109/ESSCIRC.2016.7598259

Li, Y., Fan, C., Li, Y., Wu, Q., & Ming, Y. (2018). Improving deep neural network with Multiple Parametric Exponential Linear Units. *Neurocomputing*, *301*, 11–24. https://doi.org/10.1016/j.neucom.2018.01.084

Luo, L., Xiong, Y., Liu, Y., & Sun, X. (2019). *Adaptive Gradient Methods with Dynamic Bound of Learning Rate*. (2018), 1–19. Retrieved from http://arxiv.org/abs/1902.09843

Nilsson, N. J. (2005). *INTRODUCTION TO MACHINE LEARNING AN EARLY DRAFT OF A PROPOSED TEXTBOOK Department of Computer Science*.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). Dl-物体検出05-2_2015_Yolo(Cvpr). *Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)*, 779–788. https://doi.org/10.1109/CVPR.2016.91

Rie Johnson, A., & Zhang, T. (2013). Accelerating Stochastic Gradient Descent using Predictive Variance Reduction. *Minerva Ginecologica*, *20*(21), 1728–1733.

Sainath, T. N., Kingsbury, B., Soltau, H., & Ramabhadran, B. (2013). Optimization techniques to improve training speed of deep neural networks for large speech tasks. *IEEE Transactions on Audio, Speech and Language Processing*, *21*(11), 2267–2276. https://doi.org/10.1109/TASL.2013.2284378

Sun, C., Shrivastava, A., Singh, S., & Gupta, A. (2017). Revisiting Unreasonable Effectiveness of Data in Deep Learning Era. *Proceedings of the IEEE International Conference on Computer Vision*, *2017-October*, 843–852. https://doi.org/10.1109/ICCV.2017.97

Taylor, L., & Nitschke, G. (2017). *Improving Deep Learning using Generic Data Augmentation*. Retrieved from http://arxiv.org/abs/1708.06020

Wu, S., Li, G., Deng, L., Liu, L., Wu, D., Xie, Y., & Shi, L. (2018). L1-Norm Batch Normalization for Efficient Training of Deep Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 1–8. https://doi.org/10.1109/TNNLS.2018.2876179

Wu, Z., Chen, X., Gao, Y., & Li, Y. (2018). Rapid target detection in high resolution remote sensing images using YOLO Model. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(3), 1915–1920. https://doi.org/10.5194/isprs-archives-XLII-3-1915-2018

Ye, C., Yang, Y., Fermuller, C., & Aloimonos, Y. (2017). *On the Importance of Consistency in Training Deep Neural Networks*. 1–13. Retrieved from http://arxiv.org/abs/1708.00631

Zhang, Y., Sohn, K., Villegas, R., Pan, G., & Lee, H. (2015). Improving object detection with deep convolutional networks via Bayesian optimization and structured prediction. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *07-12-June-2015*, 249–258. https://doi.org/10.1109/CVPR.2015.7298621