References

[1] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, “On Calibration of Modern Neural Networks,” arXiv:1706.04599 [cs], Aug. 2017. Available: http://arxiv.org/abs/1706.04599

[2] L. Alzubaidi et al., “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” J Big Data, vol. 8, no. 1, p. 53, Dec. 2021, doi: 10.1186/s40537-021-00444-8.

[3] G. Hinton, O. Vinyals, and J. Dean, “Distilling the Knowledge in a Neural Network,” arXiv:1503.02531 [cs, stat], Mar. 2015. Available: http://arxiv.org/abs/1503.02531

[4] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. 2016. MIT Press

[5] Platt, John et al. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Advances in large margin classifiers, 10(3): 61–74, 1999

[6] Dan Hendrycks and Kevin Gimpel. A baseline for detecting misclassified and out-of-distribution examples in neural networks. ICLR, 2017

[7] A. Niculescu-Mizil and R. Caruana, “Predicting good probabilities with supervised learning,” in Proceedings of the 22nd international conference on Machine learning - ICML ’05, Bonn, Germany, 2005, pp. 625–632. doi: 10.1145/1102351.1102430.

[8] A. S. Mozafari, H. S. Gomes, W. Leão, S. Janny, and C. Gagné, “Attended Temperature Scaling: A Practical Approach for Calibrating Deep Neural Networks,” arXiv:1810.11586 [cs, stat], May 2019.Available: http://arxiv.org/abs/1810.11586

[9] B. Ji, H. Jung, J. Yoon, K. Kim, and Y. Shin, “Bin-wise Temperature Scaling (BTS): Improvement in Confidence Calibration Performance through Simple Scaling Techniques,” arXiv:1908.11528 [cs], Sep. 201. [Online]. Available: http://arxiv.org/abs/1908.11528

[10] A. S. Mozafari, H. S. Gomes, and C. Gagne, “A Novel Unsupervised Post-Processing Calibration Method for DNNS with Robustness to Domain Shift,” arXiv:1911.11195 [cs, stat], Nov. 2019.Available: http://arxiv.org/abs/1911.11195

[11] A. Mozafari, H. Gomes, S. Janny, and C. Gagné, A New Loss Function for Temperature Scaling to have Better Calibrated Deep Networks. 2018.

[12] C. Tomani, D. Cremers, and F. Buettner, "Parameterized Temperature Scaling for Boosting the Expressive Power in Post-Hoc Uncertainty Calibration," arXiv:2102.12182 [cs], Feb. 2021, Available: http://arxiv.org/abs/2102.12182

[13] J. Nixon et al., "Measuring Calibration in Deep Learning," arXiv:1904.01685 [cs, stat], Aug. 2020, Available: http://arxiv.org/abs/1904.01685

[14] Naeini, Mahdi Pakdaman, Cooper, Gregory F, and Hauskrecht, Milos. Obtaining well calibrated probabilities using bayesian binning. In AAAI, pp. 2901, 2015.

[15] A. Niculescu-Mizil and R. Caruana, "Predicting good probabilities with supervised learning," in Proceedings of the 22nd international conference on Machine learning - ICML '05, Bonn, Germany, 2005, pp. 625–632. doi: 10.1145/1102351.1102430.

[16] LeCun, Yann, Bottou, Leon, Bengio, Yoshua, and Haffner, Patrick. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, 1998

[17] He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, and Sun,Jian. Deep residual learning for image recognition. In CVPR, pp. 770–778, 2016.

[18] Zadrozny, Bianca and Elkan, Charles. Obtaining calibrated probability estimates from decision trees and naive bayesian classifiers. In ICML, pp. 609–616, 2001.

[19] Zadrozny, Bianca and Elkan, Charles. Obtaining calibrated probability estimates from decision trees and naive bayesian classifiers. In ICML, pp. 609–616, 2001.

[20] D. Hendrycks, N. Mu, E. D. Cubuk, B. Zoph, J. Gilmer, and B. Lakshminarayanan, "AugMix: A Simple Data Processing Method to Improve Robustness and Uncertainty," arXiv:1912.02781 [cs, stat], Feb. 2020. Available: http://arxiv.org/abs/1912.02781

[21] S. Liang, Y. Li, and R. Srikant, "Enhancing The Reliability of Out-of-distribution Image Detection in Neural Networks," arXiv:1706.02690 [cs, stat], Aug. 2020. [Online]. Available: http://arxiv.org/abs/1706.02690

[22] Gabriel Pereyra, George Tucker, Jan Chorowski, Lukasz Kaiser, and Geoffrey E. Hinton. Regularizing neural networks by penalizing confident output distributions. In ICLR, Workshop Track Proceedings, 2017.

[23] Zhang, Chiyuan, Bengio, Samy, Hardt, Moritz, Recht, Benjamin, and Vinyals, Oriol. Understanding deep learning requires rethinking generalization. In ICLR, 2017.

[24] Naeini, Mahdi Pakdaman, Cooper, Gregory F, and Hauskrecht, Milos. Obtaining well calibrated probabilities using bayesian binning. In AAAI, pp. 2901, 2015.

[24] Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, and Youngjoon Yoo. CutMix: Regularization strategy to train strong classifiers with localizable features. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 6023–6032, 2019.

[25] Cubuk, E. D., Zoph, B., Mane, D., Vasudevan, V., and Le, Q. V. Autoaugment: Learning augmentation strategies from data. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 113–123, 2019.

[26] Y. Wen, D. Tran, and J. Ba. BatchEnsemble: an Alternative Approach to Efficient Ensemble and Lifelong Learning. In International Conference on Learning Representations, 2020.

[27] Lakshminarayanan, Balaji, Pritzel, Alexander, and Blundell, Charles. Simple and scalable predictive uncertainty estimation using deep ensembles. arXiv preprint arXiv:1612.01474, 2016

[28] Hara, K., Saitoh, D. & Shouno, H. Analysis of dropout learning regarded as ensemble learning. In Proc. 25th Int. Conf. Artificial Neural Networks 72–79 (ICANN, 2016).

[29] F. Laakom, J. Raitoharju, A. Iosifidis, J. Nikkanen, and M. Gabbouj, “Monte carlo dropout ensembles for robust illumination estimation,” in 2021 International Joint Conference on Neural Networks (IJCNN). IEEE, 2021, pp. 1–7.

[30] Y. Wen, G. Jerfel, R. Muller, M. W. Dusenberry, J. Snoek, B. Lakshminarayanan, and D. Tran, “Combining ensembles and data augmentation can harm your calibration,” in International Conference on Learning Representations, 2021

[31] Rafael Müller, Simon Kornblith, and Geoffrey Hinton. When does label smoothing help? In Advances in Neural Information Processing Systems, 2019

[32] M. Minderer, J. Djolonga, R. Romijnders, F. Hubis, X. Zhai, N. Houlsby, D. Tran, and M. Lucic. Revisiting the calibration of modern neural networks. arXiv preprint arXiv:2106.07998, 2021.

[33] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998a). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86, 2278–2324.

[34] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In ICLR, 2015.

[35] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[36] G. Huang, Z. Liu, K. Q. Weinberger, and L. Maaten. Densely connected convolutional networks. In CVPR, 2017.

[37] L. Frenkel and J. Goldberger. Network Calibration by Class-based Temperature Scaling. In European Signal Processing Conference (EUSIPCO) , 2021.

[38] Krizhevsky, Alex and Hinton, Geoffrey. Learning multiple layers of features from tiny images. Technical report, University of Toronto, 2009.