













- [9] A. Bibi, H. Abidi and O. Dhaouadi, "SeqCapsGAN: Generating Stylized Image Captions," 2020. Accessed on: November, 19, 2020, [Online].
- [10] S. Sabour, N. Frosst and G. E. Hinton, "Dynamic Routing Between Capsules," in Proc. NIPS, Long Beach, CA, USA, 2017, pp. 3859-3869.
- [11] A. Shah, E. Kadam, H. Shah, S. Shinde and S. Shingade, "Deep Residual Networks with Exponential Linear Unit," in Proc. VisionNet, Jaipur, RJ, India, 2016, pp. 59-65.
- [12] T. Dozat. (May, 2016). Incorporating Nesterov Momentum Into Adam. Presented at ICLR. [Online]. Available: <https://openreview.net/pdf/OM0jvwB8jlp57ZJjtNEZ.pdf>
- [13] P. Zhou, J. Feng, C. Ma, C. Xiong, S. HOI and W. E., "Towards Theoretically Understanding Why SGD Generalizes Better Than Adam in Deep Learning," in Proc. NEURIPS, 2020, pp. 21285-21296.
- [14] D. Masters and C. Luschi, "Revisiting Small Batch Training for Deep Neural Networks," arXiv preprint arXiv:1804.07612, 2018. Accessed on: January, 8, 2021, [Online].
- [15] P. M. Radiuk, "Impact of Training Set Batch Size on the Performance of Convolutional Neural Networks for Diverse Datasets," *Inf. Tech. and Mgmt. Sci.*, vol. 20, pp. 20-24, Dec. 2017. Accessed on: December, 21, 2021, DOI: 10.1515/itms-2017-0003, [Online].
- [16] Q. Fu, Y. Liu and Z. Xie, "EECS442 Final Project Report," Univ. Michigan, Ann Arbor, MI, USA, Rep. WI2020, 2019.
- [17] M. Maire, S. Belongie, B. Lubomir, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, P. Dollar and T. Y. Lin, "Microsoft COCO: Common Objects in Context," in Proc. ECCV, Zurich, ZU, Switzerland, 2014, pp. 740-755.
- [18] M. Arjovsky, S. Chintala and L. Bottou, "Wasserstein GAN," in Proc. ICML, Sydney, NSW, Australia, 2017, pp. 214-223.
- [19] C. Lin, "Rouge: A package for automatic evaluation of summaries," in Proc. ACL Workshop, Barcelona, Spain, 2004, pp. 78-81.
- [20] R. Vedantam, C. Lawrence Zitnick and D. Parikh, "Cider: Consensus-based image description," in Proc. IEEE CVPR, Boston, MA, USA, 2015, p. 4566-4575.