GTID: 90356732

# An Improvement on Bag of Tricks for Image Classification with Convolutional Neural Networks

Vivek Poovathoor
vpoovathoor3@gatech.edu

## Project summary

The work done in [1] showed that successive adjustments to the training process of a Convolutional Neural Net (CNN) lead to significant improvements in model accuracy. Such pre-processing and training adjustments include increasing the batch size, linear scaling of the learning rate, cosine learning rate decays to name a few. Additionally, [1] shows that specific model architecture tweaks lead to as much as a 1% improvement (compared to baseline ResNet-50) in validation accuracy (with a cost of slightly longer training time).

This research project will attempt to improve the validation accuracy and training time results of the ResNet architecture. The first benchmark will be to replicate the results from [1]. Then this project will explore the effects of using a different optimization function other than mini-batch stochastic gradient descent (SGD) and Nesterov Accelerated Gradient (NAG) descent, e.g. RMSProp, Adam optimzer, etc. In addition, different ResNet architectures (not mentioned in [1]) will be evaluated. This project will also identify and implement pre-processing techniques not used in [1] to improve validation accuracy on the same dataset. Together, with an improved ResNet architecture and different pre-processing techniques, this project will determine if the validation accuracy and timing results significantly improve.

And finally this project will also attempt to generalize and evaluate the results of the aforementioned modifications on a VGG architecture and Inception-V3 to showcase if the findings generalize well.

## Projected Obstacles/Challenges

1. Creating scripts that run efficiently on large datasets like ImageNet (to reduce debugging-iterations).
2. Successfully replicating the accuracy and timing results of baseline ResNet architecture (working with ResNet architecture for the first time).
3. Promptly identifying and understanding why certain learning rate schedules work better than others in ResNet architectures.
4. Choosing a ResNet architecture that can outperform the variations found in [1] and proving that it performs better.
5. Finding original pre-processing techniques that improve those mentioned in [1] on a large dataset such as ImageNet.

## Tentative Timeline of Tasks

1. Download and Pre-Process CIFAR-10, CIFAR-100 and ImageNet datasets (Oct. 7, 2020).
2. Implement a baseline ResNet-50 CNN model based on Algorithm 1 in [1] (Oct. 17, 2020).
3. Implement variations of ResNet-50 model from [1]. and identify additional model architectures (Oct. 22, 2020).
4. Begin implementing pre-processing techniques in [1] and additonal techniques (Oct. 30, 2020).
5. Implement and test additional pre-processing techniques and compare results with those found in [1] (Nov. 3, 2020).
6. Finish research on alternative ResNet-50 architecture tweaks and start training (Nov. 11, 2020).
7. Identify and start training best performing ResNet architecture variation with most best pre-processing techniques (Nov. 23, 2020).
8. Run training on VGG and Inception-V3 nets with chosen pre-processing techniques (Nov. 27, 2020)

## References


1. T. He., Z. Zhang, H. Zhang, Z. Zhang, J. Xie and M. Li, "Bag of Tricks for Image Classification with Convolutional Neural Networks," 2018 *arXiv*, 1812.01187
2. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.
3. P. Goyal, P. Dollar, R. B. Girshick, P. Noordhuis, L. Wesolowski, A. Kyrola, A. Tulloch, Y. Jia, and K. He. Accurate, large minibatch SGD: training imagenet in 1 hour. CoRR, abs/1706.02677, 2017.
4. X. Jia, S. Song, W. He, Y. Wang, H. Rong, F. Zhou, L. Xie, Z. Guo, Y. Yang, L. Yu, et al. Highly scalable deep learning training system with mixed-precision: Training imagenet in four minutes. arXiv preprint arXiv:1807.11205, 2018.
5. A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012.
6. Pawar, Dipti. "Improving Performance of Convolutional Neural Network!" Medium, Medium, 14 Aug. 2018, medium.com/@dipti.rohan.pawar/improving-performance-of-convolutional-neural-network-2ecfe0207de7.
7. Dwivedi, Priya. "Understanding and Coding a ResNet in Keras." Medium, Towards Data Science, 27 Mar. 2019, towardsdatascience.com/understanding-and-coding-a-resnet-in-keras-446d7ff84d33.
8. Fung, Vincent. "An Overview of ResNet and Its Variants." Medium, Towards Data Science, 17 July 2017, towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035.
9. Anwar, Aqeel. "Difference between AlexNet, VGGNet, ResNet and Inception." Medium, Towards Data Science, 13 June 2020, towardsdatascience.com/the-w3h-of-alexnet-vggnet-resnet-and-inception-7baaaecccc96.