Deep Learning Sans Multiplication

Quinlan Bock - A20492935 and Jane Downer - A20452471

Department of Computer Science
Illinois Institute of Technology
CS 577-02
Gady Agam

October 12, 2022

1 Problem Statement

Multiplications in neural networks are computationally expensive and slow the training process. This makes it difficult to train deep and complex networks, particularly with CNNs, and prevents improvements in efficiency. Multiplication of tensors in neural networks makes it impractical not to train on GPU or TPU; these devices are not as ubiquitous as other chip architectures, are power intensive, and expensive. AdderNet thus relieves the dependence on such devices and opens up opportunities for applications in edge devices. This problem is stated and addressed in *AdderNet: Do We Really Need Multiplications in Deep Learning?* [1] which will serve as the main research paper which this project is based on.

2 Method

AdderNet exchanges multiplication for addition in convolutional neural networks. Rather than using convolution as a measure of similarity between features and filters, it uses the l1-norm, and outputs the result of this calculation. These changes require additional modifications to maintain the performance of the network. First, AdderNet uses regularized, full-precision gradients in backpropagation, which reportedly helps improve convergence, particularly in the case of many parameters. Second, AdderNet tends to produce outputs with larger variance than traditional CNNs, which in turn decreases the size of updates; to combat this, AdderNet introduces an adaptive learning rate to ensure appropriately sized updates, which involves scaling learning rates according to the magnitude of a neuron's gradient.

3 Data

Our model will be trained on CIFAR10, a dataset of 60,000 annotated images representing 10 distinct categories. [2]

We will evaluate the trained model on both CIFAR10 and ImageNet, the latter of which is a dataset of 14 million annotated images representing over 20,000 categories. [3] Specifically, we will be using a subset of ImageNet provided through PyTorch, which contains 150,000 images containing 1000 categories.

4 Team Member Responsibilities

1. Code
   a. Train on CIFAR10: Quin
   b. Evaluate on CIFAR10: Jane
   c. Evaluate on ImageNet: Quin
2. Report
   a. Problem Statement: Jane

b. Proposed Solution: Jane
c. Implementation Details: Quin
d. Results and Discussion: shared

3. Presentation
   a. Problem Statement: Jane
   b. Background Material: Jane
   c. Proposed Solution: Jane
   d. Implementation Details: Quin
   e. Results and Discussion: shared

References

[1] Hanting Chen, Yunhe Wang, Chunjing Xu, Boxin Shi, Chao Xu, Qi Tian, and Chang Xu. Addernet: Do we really need multiplications in deep learning? *CVPR*, 2020.

[2] Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton. Cifar-10 (Canadian Institute for Advanced Research).

[3] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015.