Title: Improved Regularization of Convolutional Neural Networks
Lead TA: Tunhou Zhang
Student: Xian Sun (xs75)
Project type: Individual
Project execution plan:

(1) Objective

I will implement cutout and mixup methods to improve the generality of CNN models like ResNet on CIFAR. Try different depths of ResNet to observe how these regularization techniques address overfitting issues for deeper and deeper ResNets. Try to find the best set of hyperparameters where we can reach the best performance. Then transfer those two methods to other datasets like MNIST to see if it can work well. Combine those methods to see if we can find a method that can have high accuracy and strong robustness.

(2) Background

We know that deep learning does a good job on training accuracy which also leads to other problem: overfitting. It means that we cannot predict unseen data well if we cannot solve overfitting problems. There are a lot of techniques to improve the robustness and generality of CNN models such as cutout, mixup, filp, etc.

(3) Experiment Plan

First, I plan to implement cutout and mixup algorithm and apply it on ResNet-CIFAR architecture with different depths (e.g.,18, 20, 32, etc.) to observe how these regularization techniques address overfitting issues for deeper and deeper ResNets.

Second, I will try different hyperparameters to find the best set of hyper parameters. Third, I will apply the same hyper parameter on other dataset like SVHN and Fashion MNIST to see if it works well. Fourth, I will try to combine those methods to find a good balance between accuracy and generalization.

(4) Challenges

There is no guarantee that the cutout and mixup will work very well for other model with different dataset. And there is also no guarantee that with combination of those methods, I can find a good balance between accuracy and generalization. Besides, training CNN models need a lot of time. I need to start my work as soon as possible.

(5) Time schedule

First week: Implement those methods. Second week: Apply methods to ResNet-CIFAR and find the best hyperparameters. Third week: Third, I will apply the same hyper parameter on other dataset like SVHN and Fashion MNIST to see if I can find a good balance between accuracy and generalization. Fourth week: make the slides for presentation.

(6) Expected results

I will plot the figures about accuracy vs. epochs for different method or different depths of ResNet with best performace. I will also make a table to record all test accuracy on different situations I am discussing above. Make conclusions about all observations.