

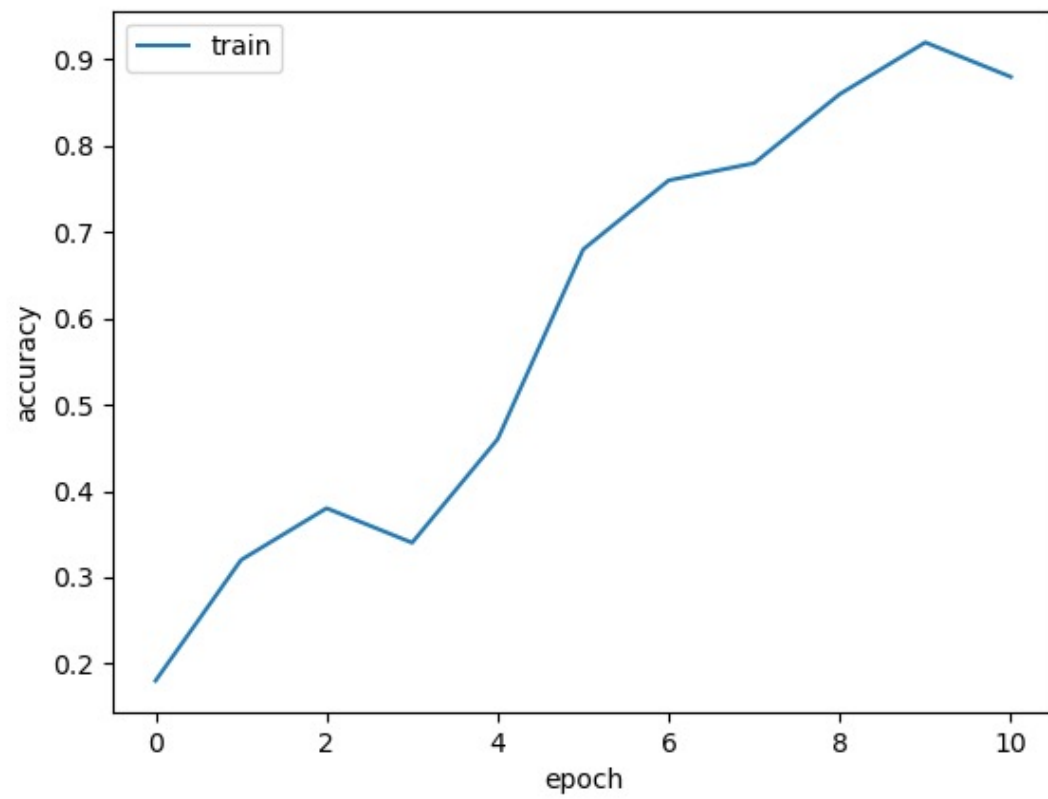
Collaborators:

Machine learning theory book, pytorch docs,
research papers given

Assignment 2 Writeup

Name: Bojun Yang
GT Email: byang301@gatech.edu
GT ID: byang301

Part-1 ConvNet



My EvalAI Model

Model description: My model was inspired by architectures for cifar10 I found in papers. All of my layers are followed by a GELU activation. It is a sequential neural net that begins with 4 sets of 2 conv layers (6 total conv) with increasing hidden layers from 128 to 512. I found that the depth of the model was more effective in increasing accuracy than width so I increased the depth instead of the width of the conv blocks. I apply max pooling, batch norm, and dropout(0.1) to each set. Then I simply have 3 fully connected layers that classify the flattened output from the conv layers. The kernel size for my conv layers was 3, stride = 1, and padding = 1.

Hyperparameter choice: I kept the hyperparameters the same as the vanilla cnn except I increased learning from 0.0001 to 0.01. I did this because I saw my accuracy was generally increasing, but never reached equilibrium. I decreased the size of the kernel from 7 to 3 because the images were already 32x32 and so a smaller kernel would be able to better filter out details from an already small image. After figuring out more optimal hyperparameters, I increased the number of epochs to get more training in and gradually increased the number of hidden nodes of each layer.

EvalAI acc: My final accuracy on EvalAI was 0.91, under the username byang301.

Data Wrangling

What's your result of training with regular CE loss on imbalanced CIFAR-10?

Fill in your per-class accuracy in the table

	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
CE Loss	.833	.8060	.2900	.1770	.0110	.0020	0	0	0	0

Put your results of CE loss and CB-Focal Loss(best) together:

	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
CE Loss	.833	.8060	.2900	.1770	.0110	.0020	0	0	0	0
CB-Focal	.5720	.1500	.1700	.3430	0	0	0	0	0	0

I observed that CB focal loss was able to achieve higher accuracies on classes with less labels in the data.