

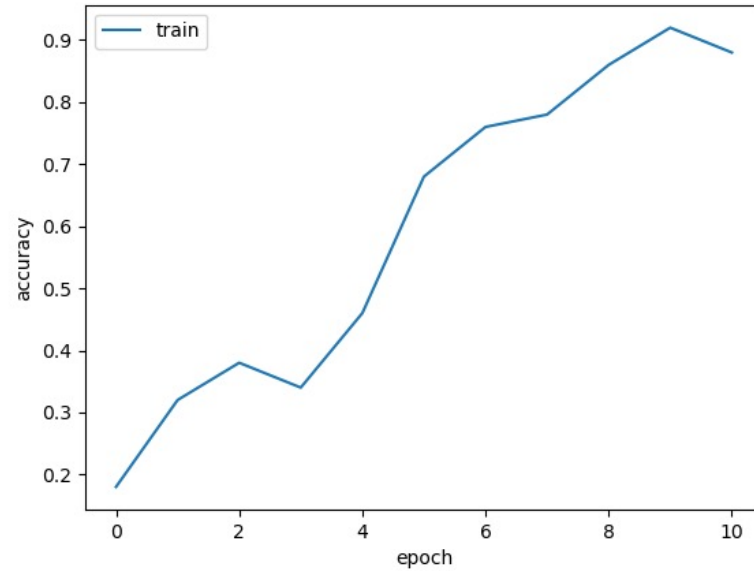
Assignment 2 Writeup

Name:

GT Email:

Part-1 ConvNet

Put your learning curve here:



My CNN Model

Describe your model design in plain text here:

My model is like VGG model except that I used less numbers of blocks. I used 3 blocks to gradually bring the number of channels from 3 to 32 to 64 to 128, while reduce the size of the image from 32x32 to 16x16 to 8x8. I used two convolution layer in each block, one to bring up the number of channels and one to maintain the number of channels (just like VGG). After each conv layer I put a relu layer and a BN layer. At the end of each block, I used a maxpooling layer with 2x2 kernel to downsample the image and a dropout layer to regularize the model. Finally, I used two FC layer to compute the class logits from flattened images obtained from the conv blocks.

Describe your choice of hyper-parameters:

I used $lr = 0.01$, $epoch = 100$, $warmup = 5$ and $steps = [90, 95]$. I have tried $lr = 0.002$ and 0.001 and found out that the learning rate is too slow. The reason why I used $steps = [90, 95]$ is that I found out that the performance is not so good after even 100 epochs ($< 90\%$ accuracy), so I think it is not meaningful to fine search the space to avoid too large steps.

What's your final accuracy on validation set?

88.31%

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	0.9850	0.9600	0.8060	0.6400	0.6210	0.4120	0.4540	0.3330	0.0920	0.0860

What's your result of training with CB-Focal loss on imbalanced CIFAR-10?

Tune the hyper-parameter beta and 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
beta=0.9999	0.9000	0.9560	0.5870	0.4680	0.5780	0.6360	0.6710	0.4940	0.4620	0.5510
beta=0.9999 9	0.8360	0.9350	0.5060	0.4700	0.6130	0.5390	0.6610	0.6490	0.4090	0.6480

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	0.9850	0.9600	0.8060	0.6400	0.6210	0.4120	0.4540	0.3330	0.0920	0.0860
CB-Focal	0.9000	0.9560	0.5870	0.4680	0.5780	0.6360	0.6710	0.4940	0.4620	0.5510

Describe and explain your observation on the result:

1. By using class balanced focal loss, classification of classes with less samples have better accuracies than regular CE loss. This shows the advantages of using CB based loss to gain better accuracy when the samples are not balanced across classes. The overall accuracy is also better (CB focal 0.6310; regular CE 0.5310). By down weighting the classes with more samples in computing the loss, CB focal loss can improve the overall accuracy by improving the performance over classes with few samples.
2. Also, except for the two classes (0 and 1) with the largest number of samples, performances of the resulting classifier on other classes are similar to each other. This reflects the fact that CB is useful in balancing the performance in imbalanced dataset.
3. Using CB loss function can not remove the bias in classification accuracy in imbalanced dataset. Overall, classes with the biggest number of samples still have better accuracy.