Computer Vision
Assignment 3

Name: Yash Sadhwani **NetID**: ys2913

Solution

Note: Please find the model and training file (training part of main file) at the following link:
`http://cims.nyu.edu/~ys2913/index.html`

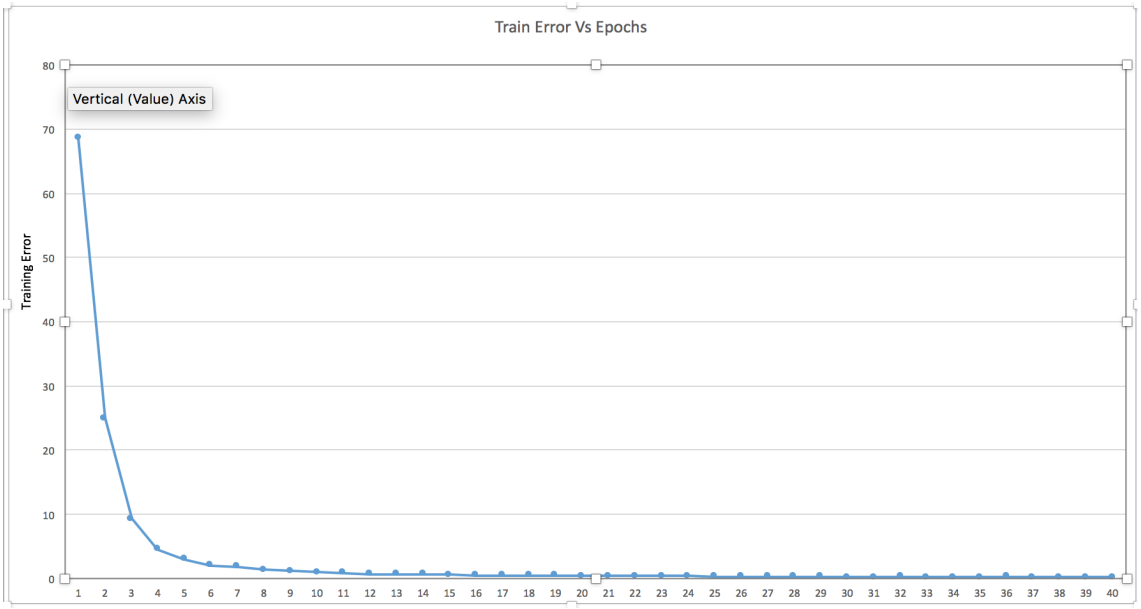
Preproceesing:

Steps involved in preprocessing Raw Data are as follows:

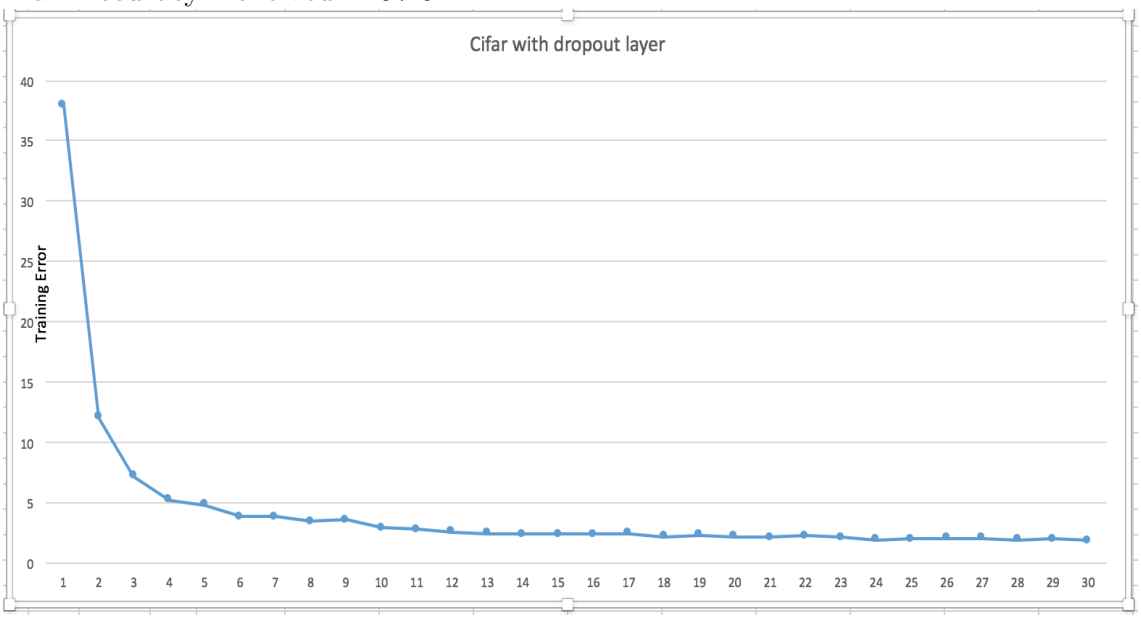
1) Cropping ROIs(region of interest) from Train and Test Images
2) Resizing the cropped image to 32x32 pixels
3) Global Normalization:
Calculating mean and standard deviation of all the training images per channel. Normalized training and test images from the mean and standard deviation calculated.
4) Converting image from RGB to YUV channel
5) Local Normalization :
Performed spatialcontrastive normalization on each image (all 3 channels).
6) Adding Jittered Data:
Adding images with translation and rotational jitters along with the original training images as training dataset.

Models Used:

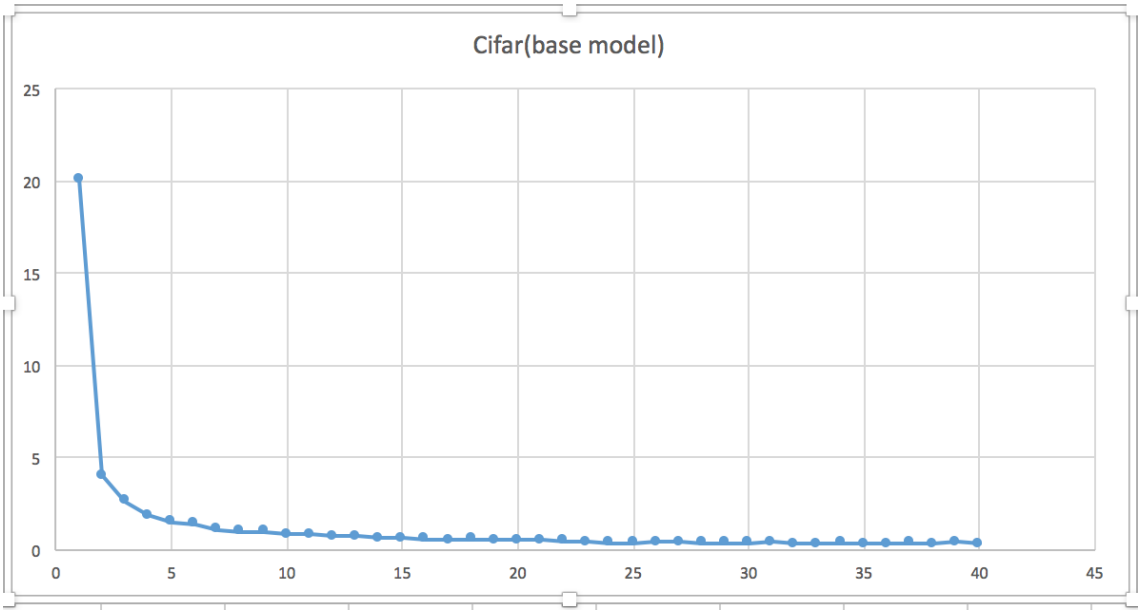
Relevant Models Used are as follows:
1) Very Deep Convolutional Network (VGG Net):
Modifications:
a) Batch Normalization of Weights
b) Addition of Dropout layer
Max Accuracy Acheived = 98.781 (submitted model)



2) Modified Cifar:
Modifications:
a) Addition of Dropout layer
Max Accuracy Acheived = 97.6



3) Cifar:
Max Accuracy Acheived = 97.2



GPU and Parallel Iterator:
Implemented changes for the code to run on GPU
Implemented parallel iterator (8 threads) for faster processing of training data. Acheived approximate 3x speedup

Experiments:
Splitting of Training Data into Train and Validation data
The Input data was split into training and validation data in the ratio of 9:1 per class
The Training data consists of 90 percent of actual training data per class
The Validation data consists of 10 percent of the remaining actual training data per class

Experiments were performed with various combinations of preprocessing and models
The highest accuracy (98.781) acheived was with the following steps:
1) Cropping ROI
2) Resizing
3) Global Normalization
4) Converting RGB to YUV
5) Spatial Contrastive Normalization
6) Addition of Jitters (4 sets of jittered data, 1 set of original data)
7) Model: Modified VGG network

Upsampling of images:
Since the distribution of training examples per class was skewed, tried to upscale the training data, by randomly appending training images so that we get the initial training data ratio of 1:1 for each class
The above mentioned experiment with the upscaled training data gave the accuracy of :98.57
Other experiments were done by combinations of removing jitters, removing cropping, Upsampling Training data, using cifar and modified cifar models.

FLOPS Calculation:

VGG Network(19 layers):

vgg:add(nn.SpatialConvolution(3, 64, 3,3, 1,1, 1,1))
Flops=3*64*32*32*3*3=1769472
vgg:add(nn.SpatialConvolution(64, 64, 3,3, 1,1, 1,1))
Flops=64*64*32*32*3*3=37748736
vgg:add(MaxPooling(2,2,2,2):ceil())
Flops=64*32*32=65536
vgg:add(nn.SpatialConvolution(64, 128, 3,3, 1,1, 1,1))
Flops=128*64*16*16*3*3=18874368
vgg:add(nn.SpatialConvolution(128, 128, 3,3, 1,1, 1,1))
Flops=128*128*16*16*3*3=37748736
vgg:add(MaxPooling(2,2,2,2):ceil())
Flops=128*16*16=32768
vgg:add(nn.SpatialConvolution(128, 256, 3,3, 1,1, 1,1))
Flops=128*256*8*8*3*3=18874368
vgg:add(nn.SpatialConvolution(256, 256, 3,3, 1,1, 1,1))
Flops=128*256*8*8*3*3=37748736
vgg:add(nn.SpatialConvolution(256, 256, 3,3, 1,1, 1,1))
Flops=128*256*8*8*3*3=37748736
vgg:add(MaxPooling(2,2,2,2):ceil())
Flops=256*8*8=16384
vgg:add(nn.SpatialConvolution(256, 512, 3,3, 1,1, 1,1))
Flops=256*512*4*4*3*3=18874368
vgg:add(nn.SpatialConvolution(512, 512, 3,3, 1,1, 1,1))
Flops=256*512*4*4*3*3=37748736
vgg:add(nn.SpatialConvolution(512, 512, 3,3, 1,1, 1,1))
Flops=256*512*4*4*3*3=37748736
vgg:add(MaxPooling(2,2,2,2):ceil())

Flops=512*4*4=8192
vgg:add(nn.SpatialConvolution(512, 512, 3,3, 1,1, 1,1))
Flops=512*512*2*2*3*3=9437184
vgg:add(nn.SpatialConvolution(512, 512, 3,3, 1,1, 1,1))
Flops=512*512*2*2*3*3=9437184
vgg:add(nn.SpatialConvolution(512, 512, 3,3, 1,1, 1,1))
Flops=512*512*2*2*3*3=9437184
vgg:add(MaxPooling(2,2,2,2):ceil())
Flops=512*2*2=2048

Total Flops=313321472