CONVOLUTIONAL NEURAL NETWORKS - PART1

Hyper-parameters of best working model:

1. Learning Rate = 0.01 (default eras learning rate)
2. Epochs = 30
3. Batch Size = 128
4. Validation Split = 0.1

I use keras callback ModelCheckpoint to store the model with best validation accuracy. After training is complete, I load out the model weights and use them to predict on test data.

Training Accuracy After 30 epochs = 99.24%
Validation Accuracy After 30 epochs = 71.72%
Test Accuracy on Moodle = 69.6%

CONVOLUTIONAL NEURAL NETWORKS - PART2

Models Experimented With:

1. **ResNet18**
2. **VGGNet16**
3. **VGGNet19**

ResNet18 Results:

1. Implemented ResNet18 i.e 4 initial convolutional layers, followed by 6 resent blocks (with skip connections) each having two layers, followed by two fully connected layers with a final softmax 10 output layer.
2. Test Accuracy after about 1 hour was only 80%.

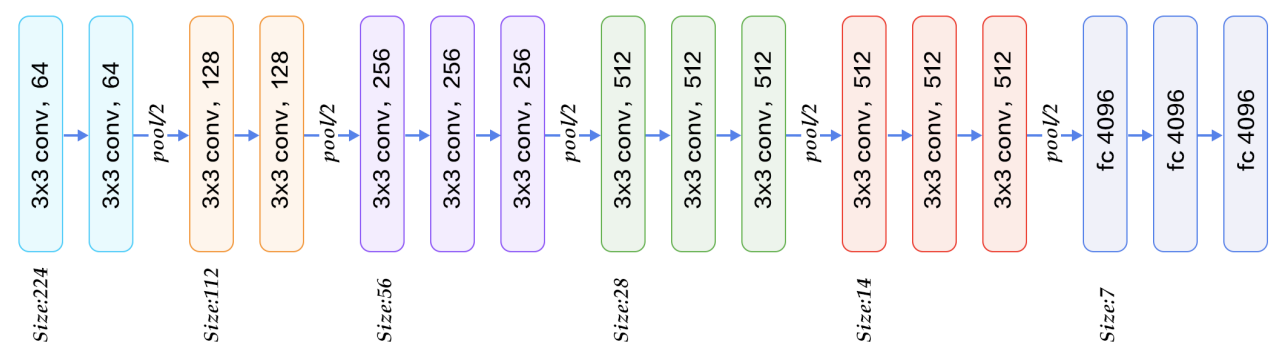
VGGNet19:

1. Architecture similar to VGG16, but with more layer
2. Too many model parameters, training happened at a very slow rate, so I abandoned this idea as well

VGGNet16:

Gives me the best accuracy.

The general architecture is:



1. It has 13 convolutional layers, followed by 3 fully connected layers. I have modified my network so that instead of the last three fully connected layers, I have only one Dense Layer of 512 size, followed by the standard Dense layer with softmax output for the ten classes.
2. I have used ONLY dropout for regularization. I have used 0.45 dropout in the first block, followed by 0.55 dropout in the remaining blocks. I have not used any L2 regularisation
3. Learning Rate = 0.01
4. Batch Size = 128
5. Epochs = 40
6. Data Augmentation: Flipping images horizontally using eras augmentation library. This does flipping with some probability during training. I am able to get a 2% improvement in test accuracy with this flipping augmentation.

Fine Tuning of Hyper-parameters:

1. Dropout: Initially I used L2 regulariser with no dropout, I found that training accuracy after 30 epochs was staling at around 75% and rising very slowly. I tried setting L2 weight decay appropriately, but to no avail. Thus I abandoned the idea of L2 regularization.
2. Initially started off with 0.4 dropout everywhere, but found overfitting was resulting. I increased the dropout to 0.5 and then to 0.55. This has given me the best accuracy so fat