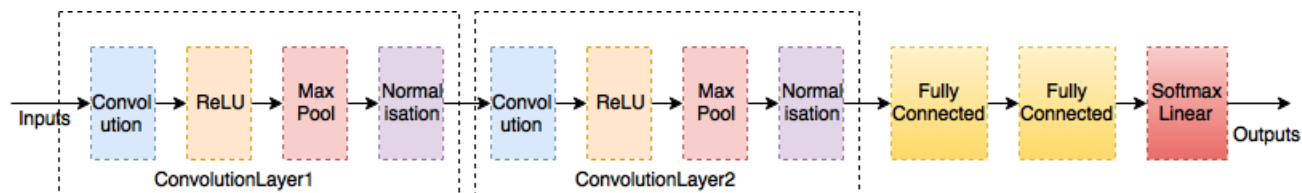


Homework Assignment 2

Sai Nikhil Maram
Advance Topics in Computer Vision

March 3, 2018

For this assignment, I have used CIFAR-10 as the dataset and built a multi-layer CNN(2 Convolution Layers, 2 fully connected layer and softmax linear layer). Best Accuracy achieved on test data is 78.7% with dropout mechanism and with 200 epochs.



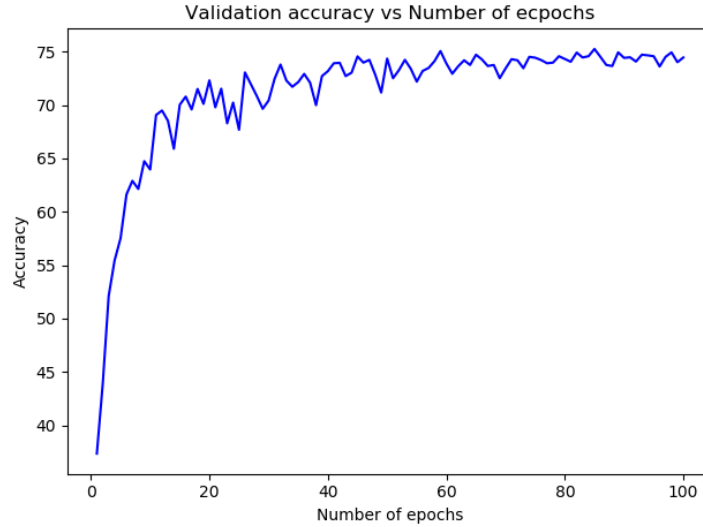
Above diagram illustrates the architecture used for CIFAR-10 classification.

1. Data is split into 45000 training, 5000 validation and 10000 test data
2. Data needs to be reshaped into [height,width,depth format] from [depth,height,width]
3. Cross entropy is considered as Loss function and ADAM optimizer is used.
4. For all the training purpose batch size is considered 128.
5. Early stopping method is employed so the data doesn't over-fit.

Different kind of experiments are carried out with the network

1. Different Initializations
 - (a) Variables are drawn from random normalization: For this initialization, the validation accuracy starts from 21% in first epoch to a maximum accuracy of 59% during training.
 - (b) Xavier Initialization: For this initialization, the validation accuracy starts from 35% in first epoch to a maximum accuracy of 74% during training.
 - (c) Variance Scaling Initialization : This gives same results as Xavier Initialization.

This clearly shows that the initialization has an effect on accuracy.



2. Different Learning Rates

- (a) initial learning rate : 0.0001, For 100 epochs the model gives an accuracy of 60%.
- (b) initial learning rate : 0.1, For 100 epochs the model gives an accuracy of 75%.

Since I have used an ADAM optimizer choosing a high learning rate gives better accuracy within limited epochs (like simulated Annealing). But when high learning rate is chosen, initial accuracies fluctuate by more which is as expected.

3. Different Normalizations

- (a) Local Response Normalization : Doesn't give better results, when considered this accuracy was only about 40%. Takes more time to compute.
- (b) Batch Normalization: Gives an accuracy of 75%.

4. Dropout mechanism was also used in fully connected layers. Although the increase in accuracy was 1%. It is observed that model converges faster when dropout mechanism is used.

5. Different Number of convolution layers. I have tried a network with 3 convolution layers also, Accuracy was increased around 1% but the training time increased by 3s for every epoch.

6. Different filter sizes and Different number of filters are tried the best combination resulted is

- (a) Convolution Layer 1: filter Size : 5, number of filters : 64
- (b) Convolution Layer 2: filter Size : 5, number of filters : 64

7. Code is evaluated on 3 platforms and time taken for an epoch in each platform is observed

- (a) 8GB RAM, 2.3GHz - 320seconds
- (b) 32GB RAM, 3.2GHz - 190seconds
- (c) K80 12GB GPU - 9 seconds