Image Classification

Harshit(1301CS19 & Sanjeevini Ganni(1301CS18)

April 17, 2017

1 Abstract of the project

Image Classification is a very important task in the field of Computer Vision. A lot of research are done worldwide to develop various image classification techniques. In this project we will used supervised deep learning for the problem. We developed a 7 layer convolutional network which behaves like other classification techniques with the resultant accuracy of 62%. Highest accuracy achieved for CIFAR 100 dataset is 75.72% by training the network with Exponential Linear Units. Our network trains well in 25-30 epochs with batch size of 32.

2 Introduction

Image Classification is the task of assigning an input image with a label from a fixed set of categories. This is one of the core problems in Computer Vision, that despite its simplicity, has a large variety of practical applications. Moreover, many other seemingly distinct Computer Vision tasks (such as object detection, segmentation) can be rediced to image classification.

2.1 Literature survey

- The authors in [1], talk about the use of convolutional neural networks for classification. The model was trained for CIFAR 10 dataset with 5 epochs and 20% error rate.
- In [2], a large, deep convolutional neural network was trained to classify the 1.3 million high-resolution images in the LSVRC-2010 ImageNet training set into the 1000 different classes. On the test data, top-1 and top-5 error rates of 39.7% and 18.9% have been achieved which is considerably better than the previous state-of-the-art results. The neural network, which has 60 million parameters and 500,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and two globally connected layers with a final 1000-way softmax. To make training faster, non-saturating neurons and a very efficient GPU implementation of convolutional nets was used. To reduce overfitting in the globally connected layers a new regularization method was used.

3 Resources

- Tensorflow
- Keras
- CIFAR-100
- Python

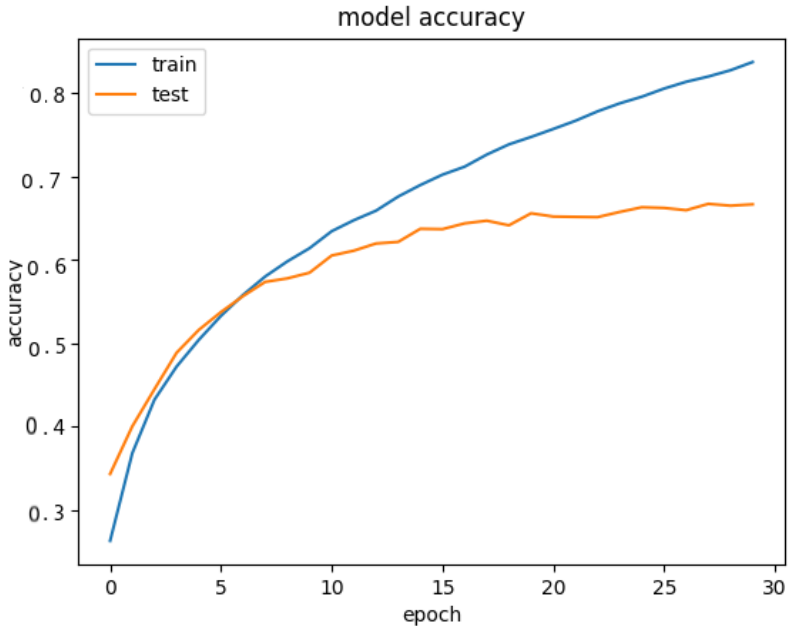

Figure 1: Accuracy vs Epochs Graph

3.1 Work done

- CIFAR-100 data cinsists of 60000 images, of 100 classes containing 600 images each. There are 500 training images and 100 test images per class. Further the 100 classe are grouped into 20 superclasses.

- We developed three netwroks for our classification problem.
  First network has accuracy around 36%. This network has a convolutional input layer with 32 feature maps with size 3*3, a rectifier activation function and a weight constraint of max norm set to 3. Then another similar layer with Max Pooling of size 2*2. Finally we have a dense fully connected layer with 100 units output layer and a softmax activation function.
  The second network has the feature layer composed of 5 convolutional layer with size 3*3 and a Max Pooling layer. The classification layer is composed of two dense layers one of size 512 and other 100. Finally we have softmax activation function for the classification. The resultant accuracy obtained for this network is 51%.
  Our third network is similar to previous one except it has some dropouts attached to convolutional layer, to avoid overfitting. This network has a accuracy around 62%. Also this network extracts HOG like features to classify the images.

- We trained the networks for training input with batch size 32 for 30 epochs. We obtain saturation without any overfitting with the help of dropouts 0.2.

- Final architecture:
  - Feature Layers : 5 Convolution layers of size 3*3, with ReLu activation, dropout of 0.2
  - Classification Layers : 1 Dense layer with size 512 and ReLu activation, dropout 0.2, final output layer with size 100 and softmax activation
  - Categorical Cross Entropy is used a the loss function.
  - Gradient descent as the optimization function with learning rate of 0.01
  - This model obtained a test accuracy of 62%.

- Code available at : https://github.com/harshit13/deep-image

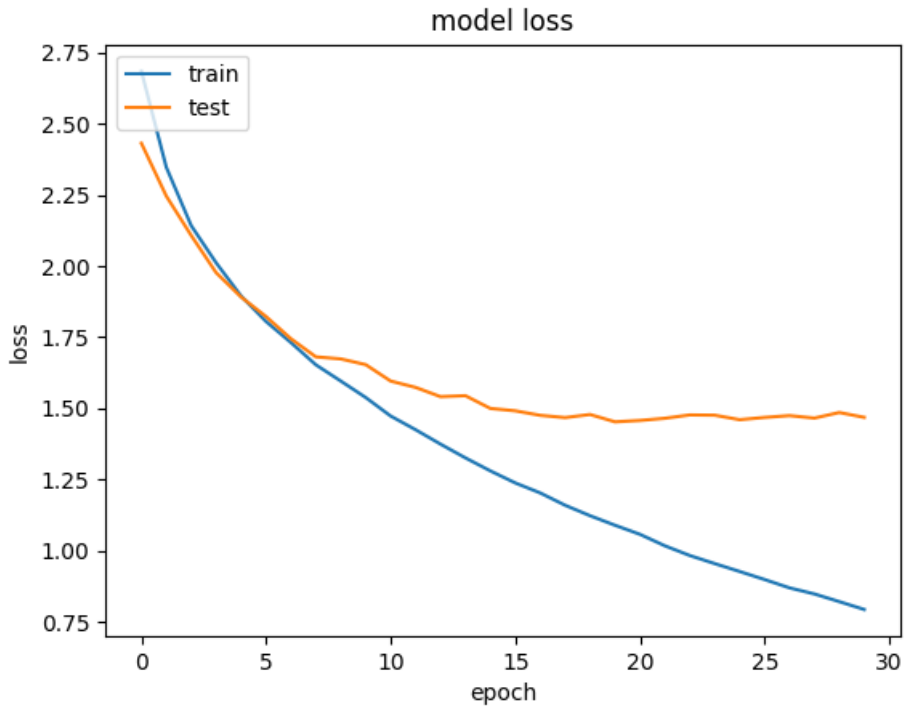

Figure 2: Model Loss vs Epochs Graph

3.2 References

- Ciresan, Dan C., et al. "Flexible, high performance convolutional neural networks for image classification." IJCAI Proceedings-International Joint Conference on Artificial Intelligence. Vol. 22. No. 1. 2011.
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.

3.3 Future work

The network seems to possess good accuracy of 62% but still, it can be improved with other optimization techniques. Also, our network trained weights can be used to classify objects in the frames of a video. A future opencv module can be developed which extracts frames from a video and applies this network over them. This would take a lot of time and processing over GPU is recommended.