CSE 574: Introduction to Machine Learning (Fall 2017)

Project 4: Introduction to Deep Learning

December 06, 2017

Group members:

1. karanhor - 50249274
2. pbisht2 - 50247429
3. rakeshsi - 50249135

Index

1. Introduction

The aim of our project is to implement a Convolutional Neural Network in order to find whether the person in the image is wearing glasses or not.

2. Dataset

We use the aligned and cropped images from the CelebA dataset.

3. Model Parameters and CNN structure

The model is a convolutional neural network (with max-pooling and dropout). The loss function used is and the optimizer is Adam. Table 1 describes the architecture of the convolutional neural network with each layer and its input parameter.

S. No.	Layers	Input Parameters
1	Input image	On reshaping 56x56x3 Other values: 72x72x3, 84x84x3
2	Convolution Layer 1	5x5x3x32
3	Activation	ReLU
4	MaxPooling	2x2
5	Convolution Layer 2	5x5x32x64
6	Activation	ReLU
7	MaxPooling	2x2
8	Fully connected layer 1	14x14x64, 1024
9	Dropout	0.5
10	Fully connected layer 2	1024x2

Table 1: Architecture of CNN

4. Experimentation and results

We performed the following steps to execute our model:

1. Extracted images and labels from the data.
2. Reduced the resolution of the original images from 178*218 pixels to 56*56, 72*72 and 84*84 pixels to reduce resolution.
3. Used different sizes of training set (9000, 13500, 18000) and testing sets (1000, 1500, 2000) respectively.
4. Applied dropout = 0.5
5. We changed the no. of filters in convolutional layers and selected the best values for our model by analysing the testing accuracy for different values.
6. Retraining the model using higher resolutions improved the accuracy of the model however, it took longer time to compute while training and testing.
7. Using bigger sizes of the training set improves accuracy to certain extent after that the model overfits resulting in degraded accuracy for testing data.
8. Using higher resolution of images using data augmentation improves the accuracy as the model gets more data from each image to analyze. However, there is a computation constraint for the resolution.

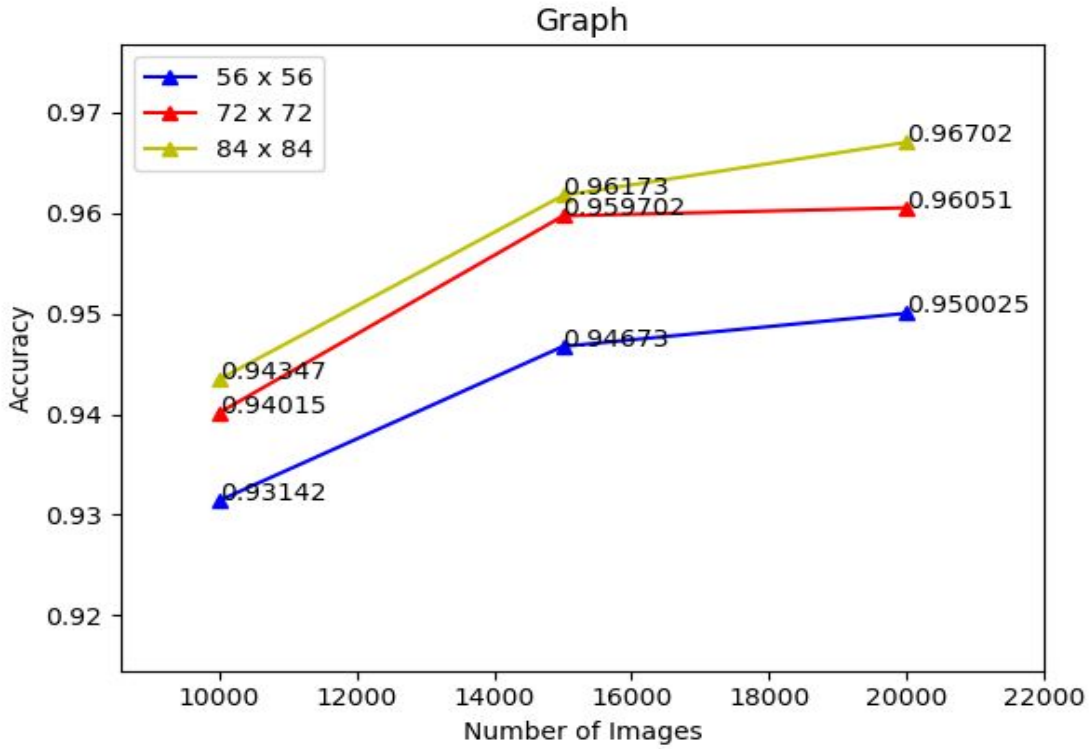

Fig.: Accuracy vs Data set size(no. of images) with different resolutions

Above graph shows different data sets of training and testing images with different resolutions of the images. We found that the as we increase the resolution of the image, the predictions are more accurate. Also the increase in the number of training images improves the accuracy upto a certain limit. However, there is a computation limit to resolution that we can adhere to while our convolutional neural network learns from data. This is because more the number of pixels means greater number of data point for each input with increased computation in each layer of the model.

5. Conclusion

We got the best accuracy from our CNN model with 84*84 pixel resolution which is 96.702% with training of 18000 images and testing 2000 images for predicting whether the celebs are whether spectacles or not.

Appendix

Data augmentation was performed but we did not see a significant increase in the accuracy in testing of the models as the dataset is already quite large and varied.

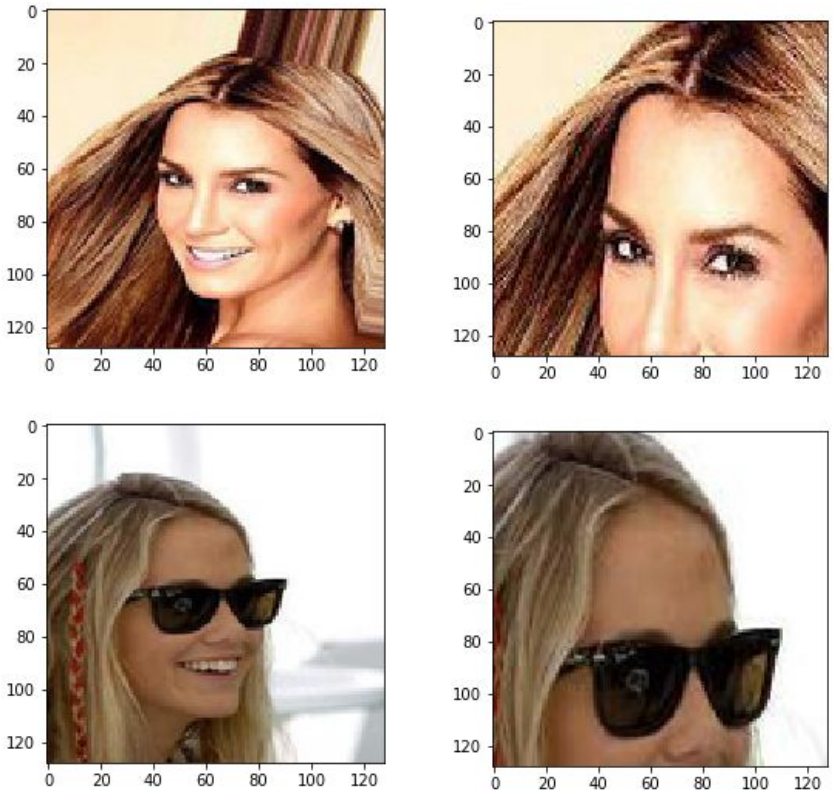

Fig: Rescaled Images(left) and cropped Images(right)

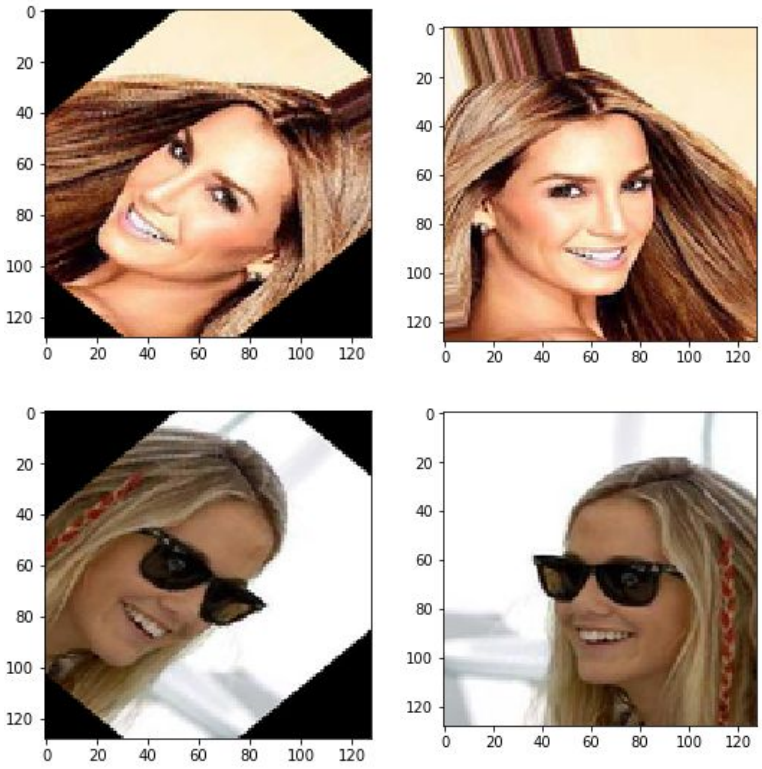

Fig: Rotated Images(left) and flipped Images(right)