UNIVERSITY OF CALGARY

ENSF 645 – Data Mining

Midterm Report

Project Title: Face Detection Using Siamese Model

Team Members:

Chetana Bijoor	**30111018**
Sanyam	**30118797**
Prabhleen Kaur	**30120534**
Neha Singh	**30109044**
Taruneesh Sachdeva	**30114398**

1. Motivation and Significance of the problem

Face Recognition has become one of the most common features used in mobile applications. Studies show that the newer generation laptops and desktops will also have this feature to ease the lives of people. Although recognizing faces is a challenging job but with the use of Siamese Networks, which are artificial neural networks provide an efficient solution for similar problem statements. Convolutional Neural Networks are useful and powerful in most cases but for this project we decided to use Siamese Neural Networks for following reasons:

• Siamese Neural Network (SNN) allows the user to choose only a handful of images from each class or from a person as compared to Convolutional Neural Networks (CNN) in which we need a large number of images belonging to different classes.

• Secondly, SNN doesn't classify images but instead learns the similarity functions which are aimed to evaluate the similarity between two objects. Some of the similarity functions correspond to some other previously explained measures, such as cosine distance, while others are statistical or probabilistic, or rely on fuzzy logic.

• Finally, on one hand, Convolutional Neural Networks are not flexible i.e., if the number of people whose faces are required change, the whole model needs to be reconstructed, Siamese Neural Networks, on the other hand, are easy to train and have a fairly simple architecture as shown in the other section of the proposal.

2. Defining dataset and pre-processing:

To create training dataset, all five team members added 5 images per person and 2 images per person were added as a test dataset. So, in total 25 images are added as the training dataset and 10 images are added as a test dataset. After defining the datasets, the images are loaded and resized for the defined dimension of (128, 128) pixels using Python Imaging Library. This results in the shape of each image as (128, 128, 3). To create an actual training dataset, the defined training images are then converted to NumPy arrays such that each image is paired with every other image and a label is assigned to each pair. Since training dataset consists of 25 images, the above-mentioned combination results in training dataset of 625 pair of images.

The label assigned to each pair of images is dependent on if the images belong to the same class or not. If both the images in a pair belongs to the same class, then its label is 1, else its label is zero. The images stored in above defined NumPy arrays are x1.npy and x2.npy files. Each of the 'x1.npy' and 'x2.npy' contains an image from each pair, such that nth image from each file, makes the nth pair with corresponding label stored at the nth position in 'y.npy' file. Thus, as a result, since we 2 sets of images, X1 and X2 of shape (625, 128, 128, 3) and 625 labels.

Based on these defined NumPy arrays and the labels, the model below defined model is trained.

3. Methodology

Model implementation methodology used to build the model is precisely described in the below image:

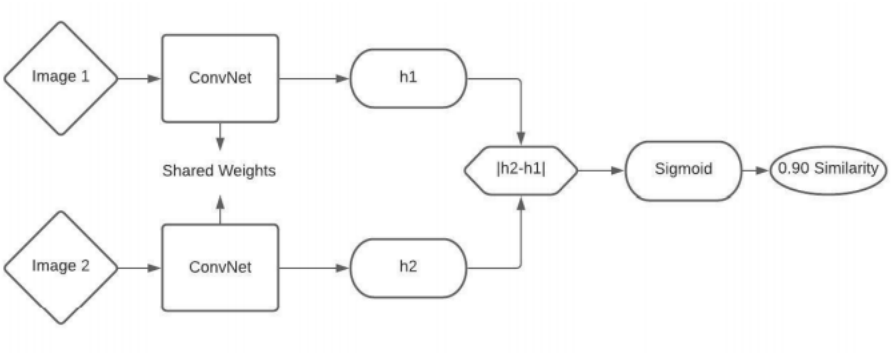

Figure 1: Siamese Model Architecture

Siamese Model for the face detection is built using the sequential model approach. Four convolution layers along with the two max pooling layers are used for building the model, with LeakyReLu as the activation function. Two instances of the model are then constructed with the exact same hyperparameters and same input layers. The outputs from the two instances of the models for each image in a pair of input images is then applied to the dense layer with 'sigmoid' as the activation function. The output of the sigmoid function is the similarity score between the two images, based on which our model learns the similarity between the two images.

The above model defined is then trained on the training dataset of images as described above. Once the model is trained, it is then tested on the test data. The test data is a set of 10 images belonging to the different classes. All the test images are passed to the trained model to calculate the similarity score. For, each image a similarity score is calculated corresponding to each class. Based on the similarity scores calculated, the image is then classified into the class, against which the highest similarity score is observed.

4. Preliminary Results

Training dataset consists of a pair of images with each pair having a label on it. If both the images in a pair belong to the same class, then its label is 1, else its label is zero. The images are stored as NumPy arrays in x1.npy and x2.npy files. Each of the 'x1.npy' and 'x2.npy' contains an image from each pair, such that nth image from each file, makes the nth pair with corresponding label stored at the nth position in 'y.npy' file.

These three files are successfully getting created on the execution of the program. Based on these files, the model is trained and tested. We tested our model for ten test images. Corresponding results shows the best confidence scores of each image along with the class into which the image is classified.

IMAGE 1 is 1 with confidence of 0.4745725393295288

IMAGE 2 is 1 with confidence of 0.49923449754714966

IMAGE 3 is 1 with confidence of 0.4350891411304474

IMAGE 4 is 3 with confidence of 0.49920007586479187

IMAGE 5 is 1 with confidence of 0.4010492265224457

IMAGE 6 is 1 with confidence of 0.4296397566795349

IMAGE 7 is 4 with confidence of 0.3761540651321411

IMAGE 8 is 2 with confidence of 0.43531787395477295

IMAGE 9 is 3 with confidence of 0.4150741398334503

IMAGE 10 is 3 with confidence of 0.49923449754714966

Actual labels for all the ten images should have been 1,1,5,5,2,2,4,4,5,5. So, based on the current results, the accuracy score of the model is less than fifty percent.

So, in the preliminary phase, we have collected our dataset and distributed it into the training and the test dataset. Further, we were able to build a basic Siamese model and test it on our dataset. The results so far achieved are written above. The accuracy so far achieved is less than 50 percent. In the data pre-processing, we have just collected the images, normalized them, and then converted into the NumPy arrays.

5. Future Steps:

In the next phase, following are the points that teams aim to focus on:

- Aim to improve the accuracy of the model
- Pre-processing of the data such that change in the dimensions of the image, removing the background of the image, and scaling of the images
- If required, we can add a greater number of the images in the training dataset
- Model can further be trained by increasing the number of epochs, and changing other hyper parameters (such as kernel size, and pool size)
- Team might also add additional convolutional layers

So, overall, the focus of the team will be increasing the accuracy of the model by changing the above-mentioned parameters and improving the pre-processing of the data. We will further focus on improving the model by changing the number of convolution layers. The team also would focus on adding the validation data set, while training the model.

GitHub Link

Below is the GitHub link of the source code files and dataset collected so far:

https://github.com/sanyam01/Enel645_FaceDetection_SiameseModel.git

6. References

- https://medium.com/@enoshshr/learning-similarity-with-siamese-neural-networks
- https://github.com/adityajn105/Face-Recognition-Siamese-Network
- https://github.com/shubham0204/Face_Recognition_with_TF
- https://medium.com/predict/face-recognition-from-scratch-using-siamese-networks-andtensorflow-df03e32f8cd0
- https://en.wikipedia.org/wiki/Siamese_neural_network
- https://towardsdatascience.com/train-validation-and-test-sets-72cb40cba9e7

7. Consensus Score

Broadly speaking, Sanyam, Taruneesh, and Neha focused on building the model and Prabhleen and Chetana focused on data collation and pre-preprocessing. Then, a main file is built which is done together as a team. So overall, the project done so far is collaborative team effort. All the team members contributed equally to the mid-term project. Hence, the consensus scores are as follows:

Name	Consensus Score
Chetana Bijoor	3
Neha Singh	3
Prabhleen Kaur	3
Taruneesh Sachdeva	3
Sanyam	3