DL Course Project Report

Siamese Neural Networks for One-shot Image Recognition

Group Members
Pradeep Kumar (B20AI029)
Pranay (B20AI030)
Utkarsh Gupta (B20ES009)

Introduction:

Face recognition is a challenging problem in computer vision that has gained significant attention in recent years. Traditional face recognition methods rely on manually engineered features, which may not be robust to variations in lighting, pose, and facial expressions. Deep learning-based approaches have shown remarkable performance in face recognition tasks, particularly in one-shot scenarios, where the goal is to recognize new images from a single example, without requiring a large amount of training data.
In this project, we aimed to implement a face recognition model using a Siamese Neural Network based on the paper by Koch et al. We present our experimental setup, methodology, and results in detail.

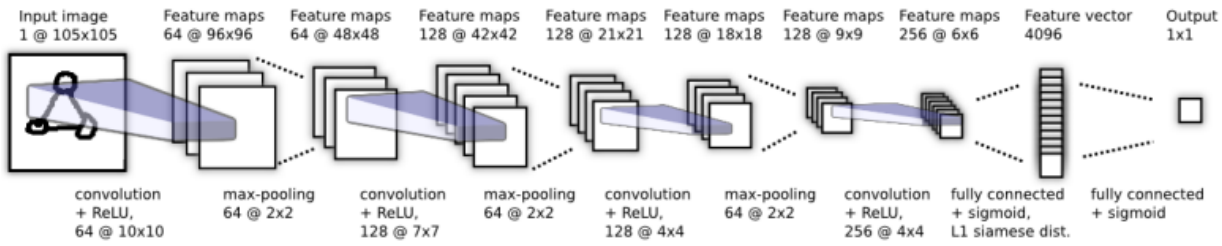
Paper Details:

The paper by Koch et al. proposes a Siamese Neural Network architecture for one-shot image recognition tasks. The network consists of two identical sub-networks that share weights and are trained to learn a similarity metric between pairs of images. During training, pairs of images are presented to the network, and the weights

are adjusted to minimize the distance between similar images and maximize the distance between dissimilar images.

Algorithm and Model Architecture Details

Siamese neural network is a type of neural network architecture that learns to differentiate between two input samples. It consists of two identical sub-networks that share the same weights, and takes in a pair of inputs. The input images are passed through a series of convolutional layers, followed by max-pooling layers and then fully connected layers. The forward_once method is used to pass a single input through the network, and the forward method is used to pass a pair of inputs through the network. The L1 distance between the embeddings of the input image and validation image is then calculated, and used to classify the pair as a match or a mismatch.

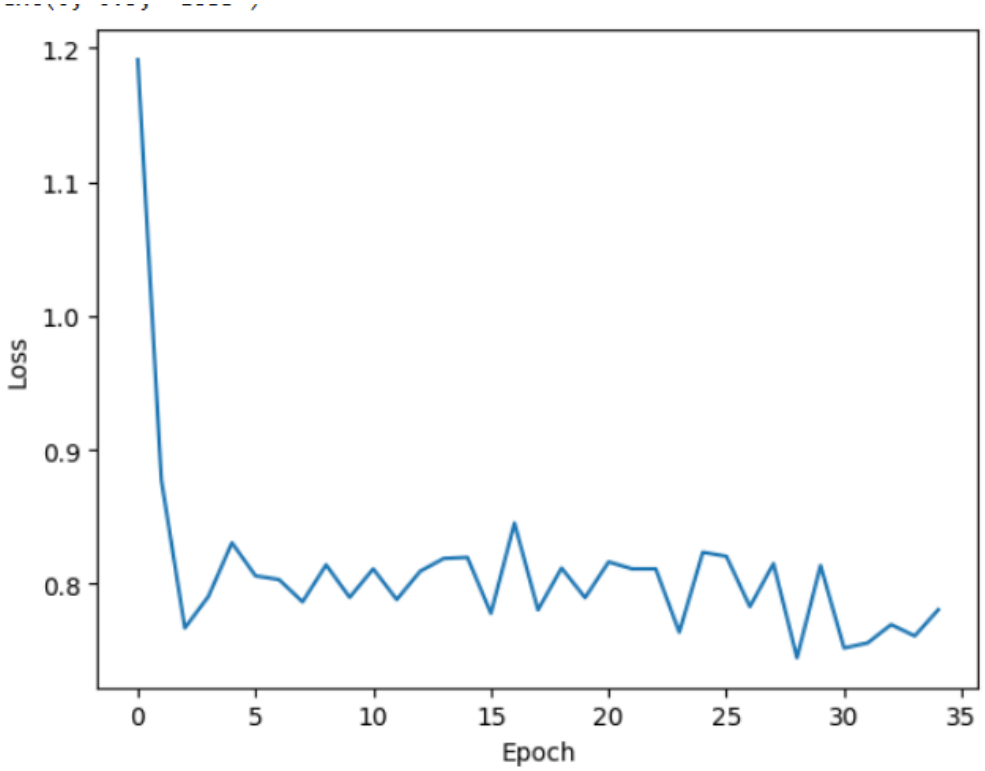


We have taken the learning rate as 0.0001.
Number of epchs as 36.
Batch size for both train and test data as 16.

Datasets Used

We used a dataset consisting of 300 positive images and 300 anchor images of a single person, as well as 300 negative images that were randomly sampled from the Labeled Faces in the Wild (LFW) dataset. The LFW dataset is a widely used dataset for face recognition tasks and contains over 13,000 face images of 5,700 people. The positive and anchor images were obtained by taking images of a single person, while the negative images were randomly sampled from the LFW dataset.

Results and Summary

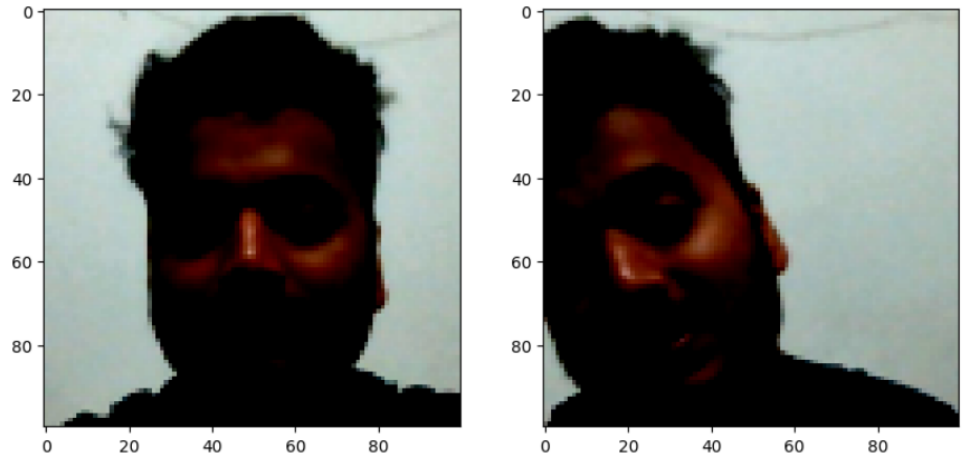



```
[ ]  predictions = siamese_model(input_image, validation_image)
```

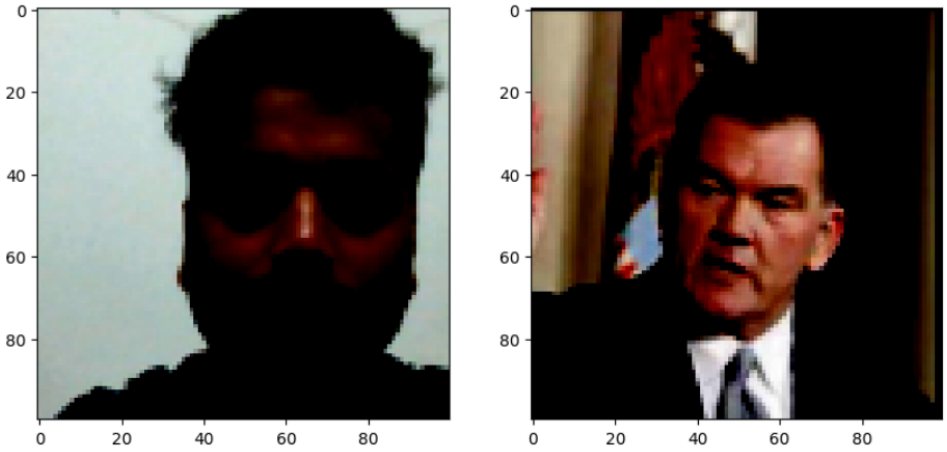
```
[ ]  # Post processing the results
     [1 if prediction > 0.5 else 0 for prediction in predictions ]

     [1, 0, 1, 1]
```

We can clearly see that after the first epoch the loss has become some what constant with slight fluctuations.

```
WARNING:matplotlib.image:Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).
WARNING:matplotlib.image:Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).
```



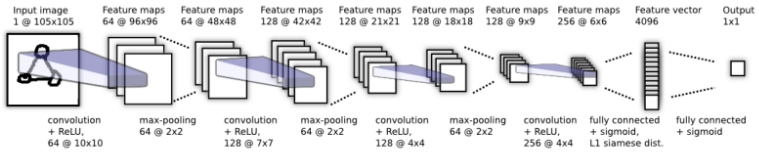
The model was able to classify the above pair of image as same.



The model was able to classify the above pair of image as different.

```
Final Recall: 1.0000, Final Precision: 0.5119.
```

Our motivation behind this project was to build a simple face recoginizer using Siamese Neural Network.
Architecture we have used for this project



The model could have been further generalized if we have considered generating more custom data of different people.

References:

[Siamese Neural Networks for One-shot Image Recognition (cmu.edu)]