Model Card - Siamese Neural Networks for One-shot Image Recognition

Model Details

- The Siamese Neural Network model consists of two identical neural networks sharing the same weights and architecture.
- The networks take two input images and generate their respective feature vectors, which are then compared using a similarity metric to determine whether the images belong to the same class or not.
- The model is trained on a few-shot learning task to learn a similarity metric that can generalize to new classes with only one or a few examples.

Intended Use

- Intended use is for the development of a facial recognition system that can recognize a person's face
- The primary intended use case is in security, facial recognition systems for phones and surveillance systems, where quick and accurate identification of individuals is crucial.

Factors

- Hardware factors such as camera resolution to collect dataset, environment factors such as lighting can affect the performance of the model

Metrics

- Loss = Binary Crossentropy.
- Optimizer = Adam
- Learning Rate = 0.0001
- Activation function = Relu in hidden layer; Sigmoid in the final layer
- Performance metrics = Precision, Recall
- Distance = L1 distance between two images i.e performing similarity calculation
- Batch Size = 32

Evaluation Data

- Validation data split consists 25 % of data
- All images are resized to 100 x 100 and normalized between 0 - 1
- Perform the validation on a set of 32 images of the same person for better prediction accuracy

Training Data

- The data is splitted into 3 different categories: positive, negative and anchor
- Positive images size = 5000, negative images size = 5000, anchor image size = 5000
- Training data split consists 75 % of data
- All images are resized to 100 x 100 and normalized between 0 - 1

Ethical Considerations

- The data collected for this work should be with the consent of the individual/people.
- Make sure that the model predictions are not biased and discriminatory against specific group like minority, skin tone or any gender
- Inaccuracies like false positive and false negative can lead to innocent people being misidentified as suspects/criminals leading to false accusations.

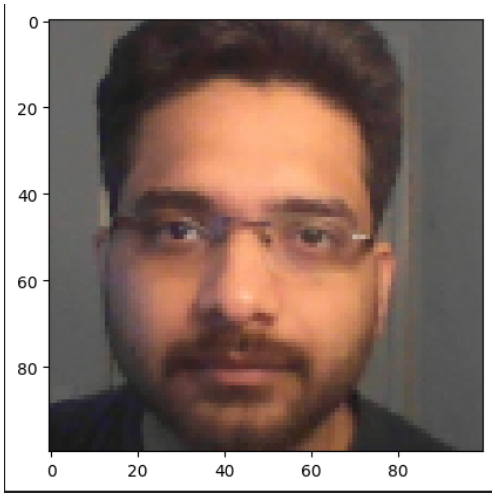
Caveats and Recommendations

- Changed resolution of images from 105 * 105 (mentioned in the research paper) to 100 * 100. (assumption that it'll not affect the model's performance)
- Collected around 5000 images for anchor and positive images.to improve the performance of the model, we need to collect more data (assumption that 5000 images are enough to train the model)
- Model is slightly overfitting because of the false positive and false negative scenarios. Can implement hyperparameter tuning, regularization, to reduce the overfitting

- This application can be scaled to support verifying multiple people instead of a single person and can also be used as a face ID app.

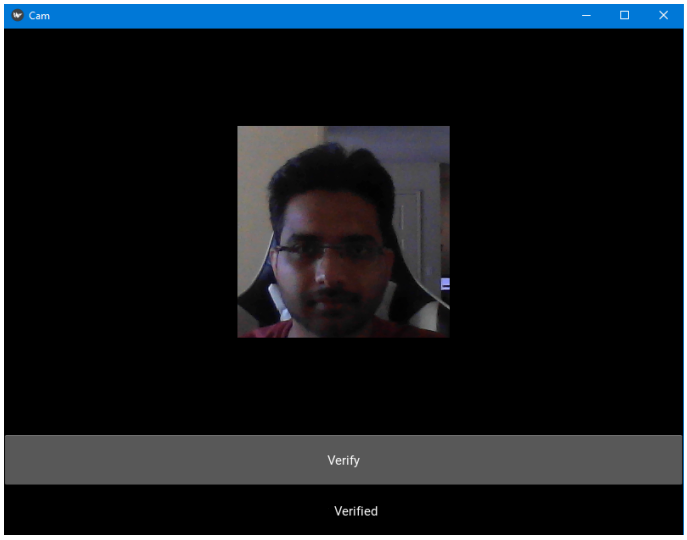
Quantitative Analyses

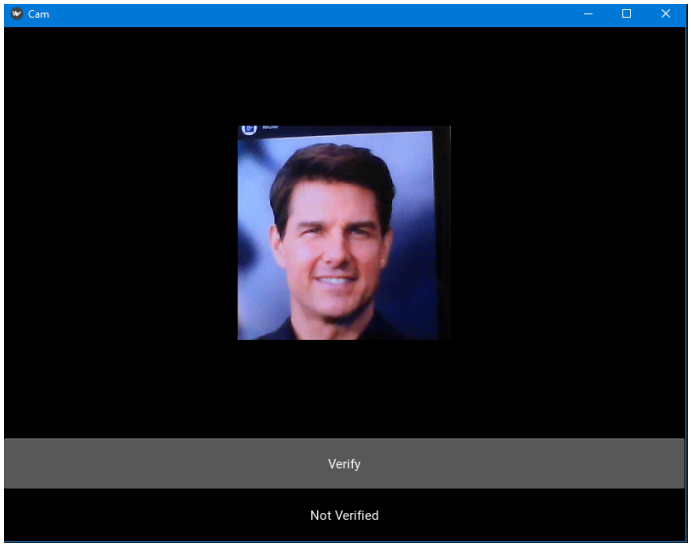
- Below is a sample image being collected during data collection phase (resolution 100 * 100)



- Performance metrics: Recall and Precision
- Training dataset: Recall = 0.98, Precision = 0.99
- Validation dataset: Recall = 1, Precision = 1

- Below is the screenshot of Verified output.



- Below is the screenshot of the not verified output . input image of Tom cruise is being compared with image of myself and the siamese model classified Tom Cruise's image as not verified
- 

- False negative case.