Detect Real Faces and Swapped Faces

—

Team Members:
Neppalli Balakrishna Prasad, Srilekha Kadambala

Introduction

Initially, we have generated swapped images using Face Swapping GAN where we have to swap the given source face on to the target face, retaining the same target pose and expression.

Source Target FSGAN

Objective

The main objective is to train a classifier using DenseNet121 Architecture to detect between the real images and swapped images generated by FSGAN.

Datasets used

- https://www.kaggle.com/ciplab/real-and-fake-face-detection
- https://www.kaggle.com/xhlulu/densenet-on-face-classification

140k real and GAN-generated faces where the real faces were taken from CelebA Dataset.

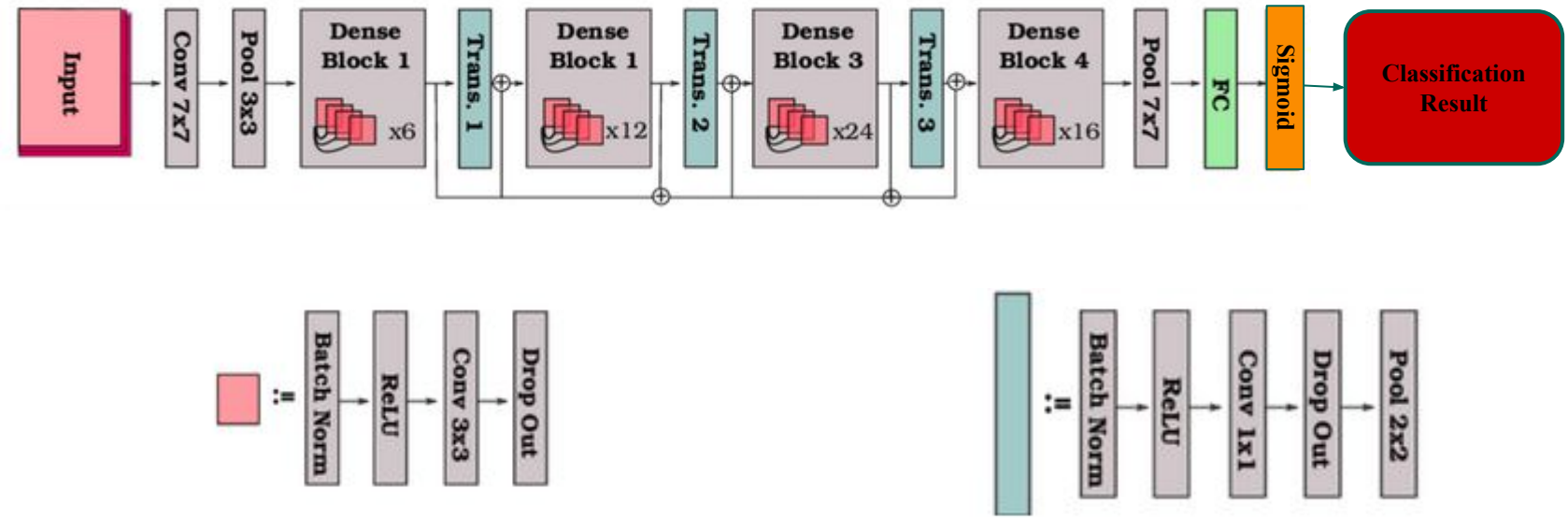
Methods Available

1. VGG
2. ResNet
3. DenseNet
4. MesoNet

Data Preprocessing

- In case of videos, split into frames.
- Resizing image as 224x224

Model Overview



Results

- 1 corresponds to fake (swapped) image, 0 corresponds to real image.
- If output is less than 0.5, then we determine it as a real image. The More nearer to 0, the more accurate result.
- If output is greater than or equal to 0.5, we determine it as fake image. More nearer to 0, the more accurate result.

Real

3.012814e-06

Fake

0.59110012

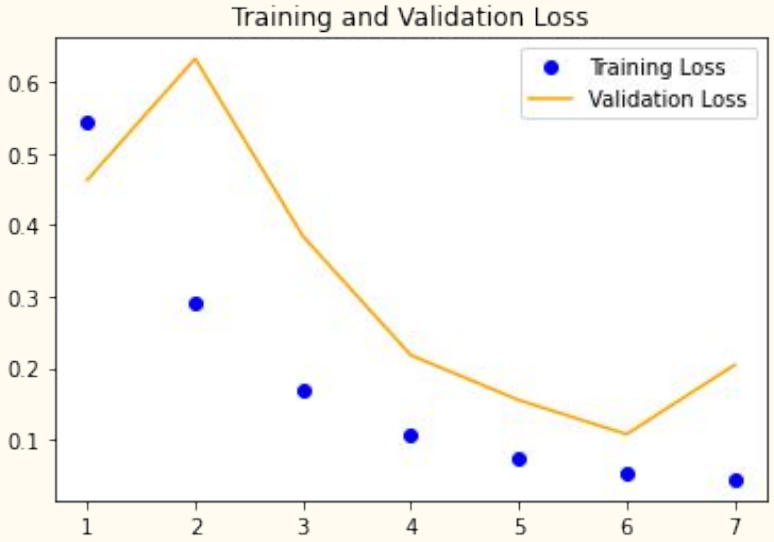
Real

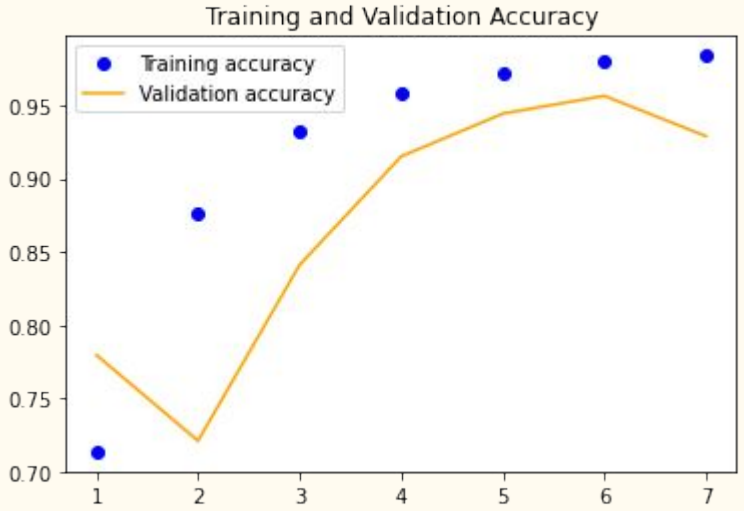
7.210045e-06

Fake

0.63654801

Results


Training and Validation Loss
Training Loss
Validation Loss
0.6
0.5
0.4
0.3
0.2
0.1
1
2
3
4
5
6
7


Training and Validation Accuracy
Training accuracy
Validation accuracy
0.95
0.90
0.85
0.80
0.75
0.70
1
2
3
4
5
6
7

Limitations

- Extreme Head Poses
- Can’t detect grayscale images
- If there are any face occlusions, then the accuracy gets low or sometimes the value shows below 0.5

Conclusion

Our results show that state-of-art CNNs are now able to distinguish with minimal mis-classification inaccuracies between fake and real data. Even with models with accuracy as high as 97% are not enough - 3% of billions of images on Google or Facebook platforms would represent an immense loss of trust from users of these interfaces.

Future Work

- Detection of gray-scale images
- Identifying faces in extreme head poses
- Training the classifier on VGG for more accurate detection

THANKYOU