Face Recognition for Mpuza

Ibrahim Momin

Overview

Problem Statement
Objective
Proposed Method
Experimental Setup
Results
Novelty
Future directions
References

Problem Statement

We want you to design and implement algorithm and model to learn from the uploaded user profile picture by checking if it is your own or a picture of other users in system you are going to use.

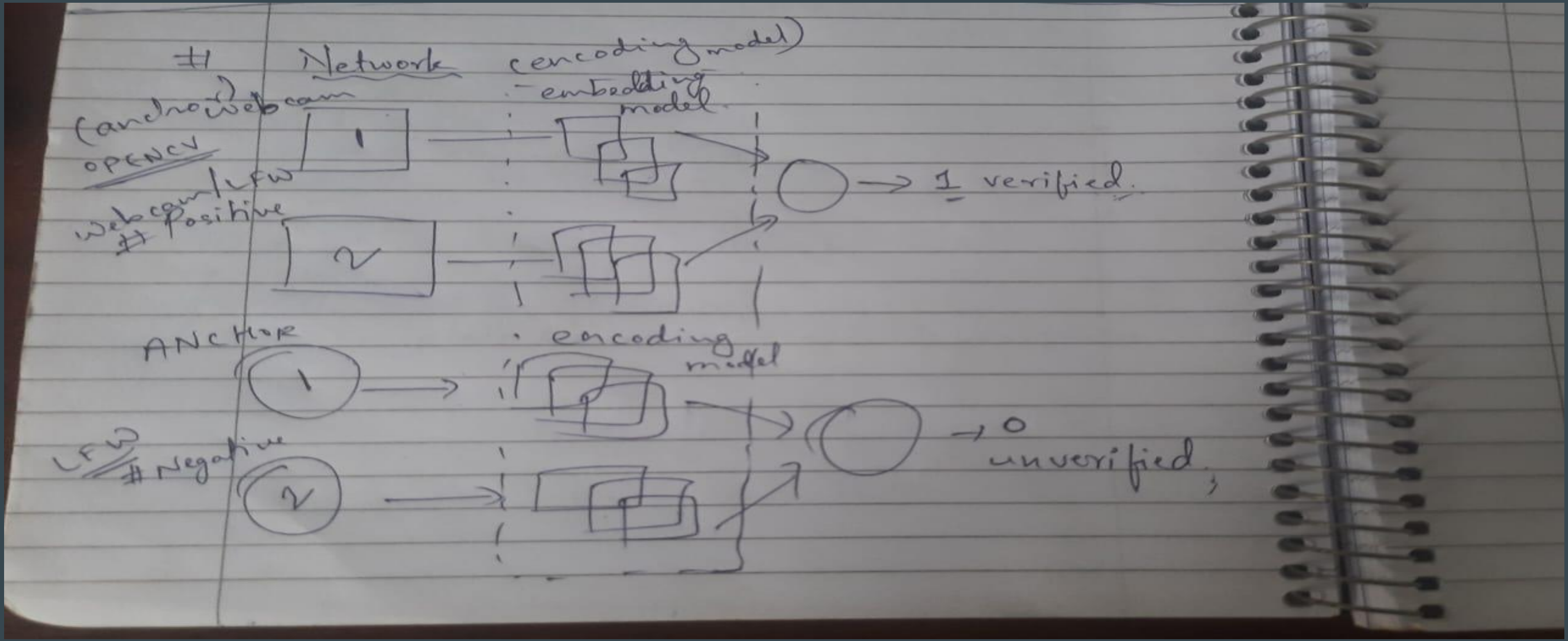
- Model can accept one input of image and check the identical one among the set of pictures exist in your dataset store.
- Increase precision and accuracy of the result up to where model can differentiate the twins
- Consider the similarity of the same person pictures within different generations like your picture if you were 12 years old and the one of today. To identify the matching.
- Consider to preserve the transformation of the face like for example to match the picture before and after accident that crush the peoples face.

Objective

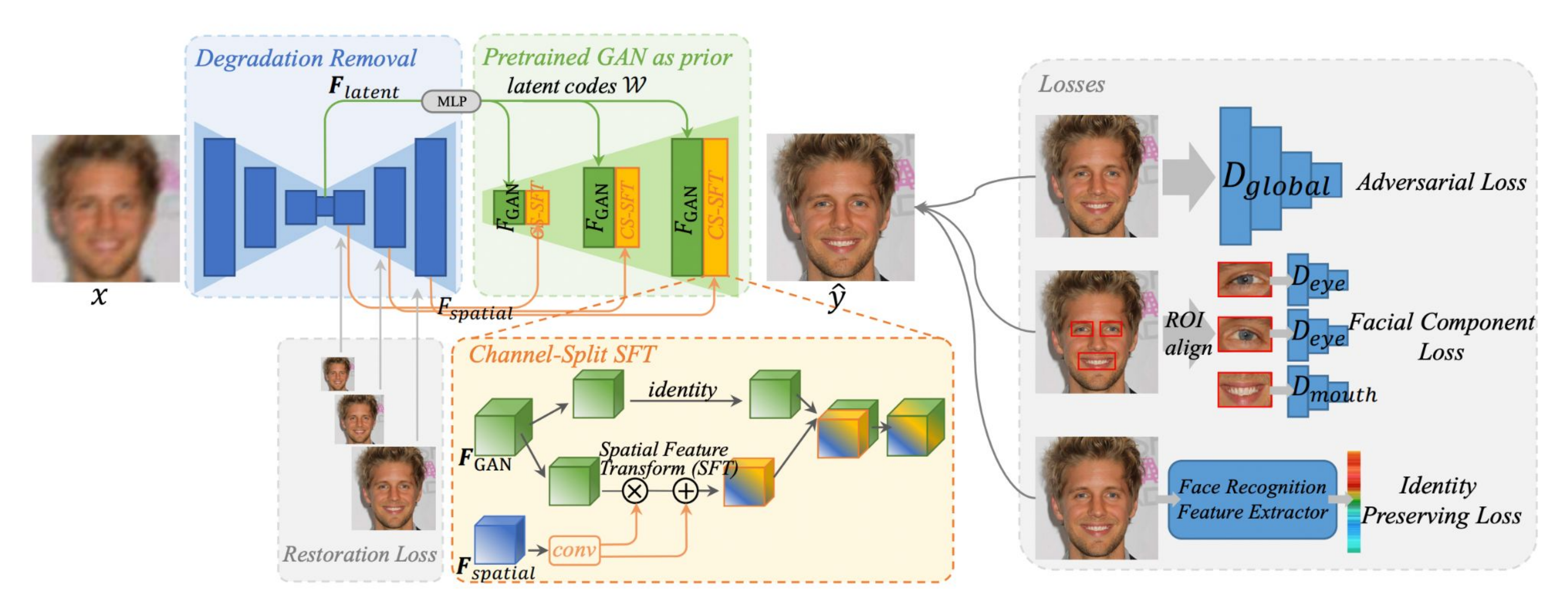
The objective of this project is to:

- Develop a Network which can perform Face Recognition while maintaining a good balance between realness and fidelity.
- Improve upon maintaining the unique identity of the person while restoring the face, so that this can be applied in realistic use-cases.
- Better understand the linking of latent code and leveraging it to do fast computations in different segments and not only Face Recognition.

Network to be followed



Approach for layers



Approach (contd.)

- **Degradation Removal**: The degradation removal module is designed to explicitly remove the above degradation and extract 'clean' features F_{latent} and $F_{spatial}$, alleviating the burden of subsequent modules.
- **Generative Facial Prior and Latent Code Mapping**: A pre-trained face GAN captures a distribution over faces in its weights of convolutions as generative prior. We leverage it to get diverse and rich facial details for our task.
- **Channel-Split Spatial Feature Transform**: In order to better preserve fidelity, we further use the input spatial features $F_{spatial}$ to modulate the GAN features F_{GAN}.

Approach (contd.)

- **Model Objectives**: The learning objective of training our GFP-GAN consists of: 1) reconstruction loss that constraints the outputs y^ close to the ground-truth y, 2) adversarial loss for restoring realistic textures, 3) proposed facial component loss to further enhance facial details, and 4) identity preserving loss.

$$L_{total} = L_{rec} + L_{adv} + L_{comp} + L_{id}$$

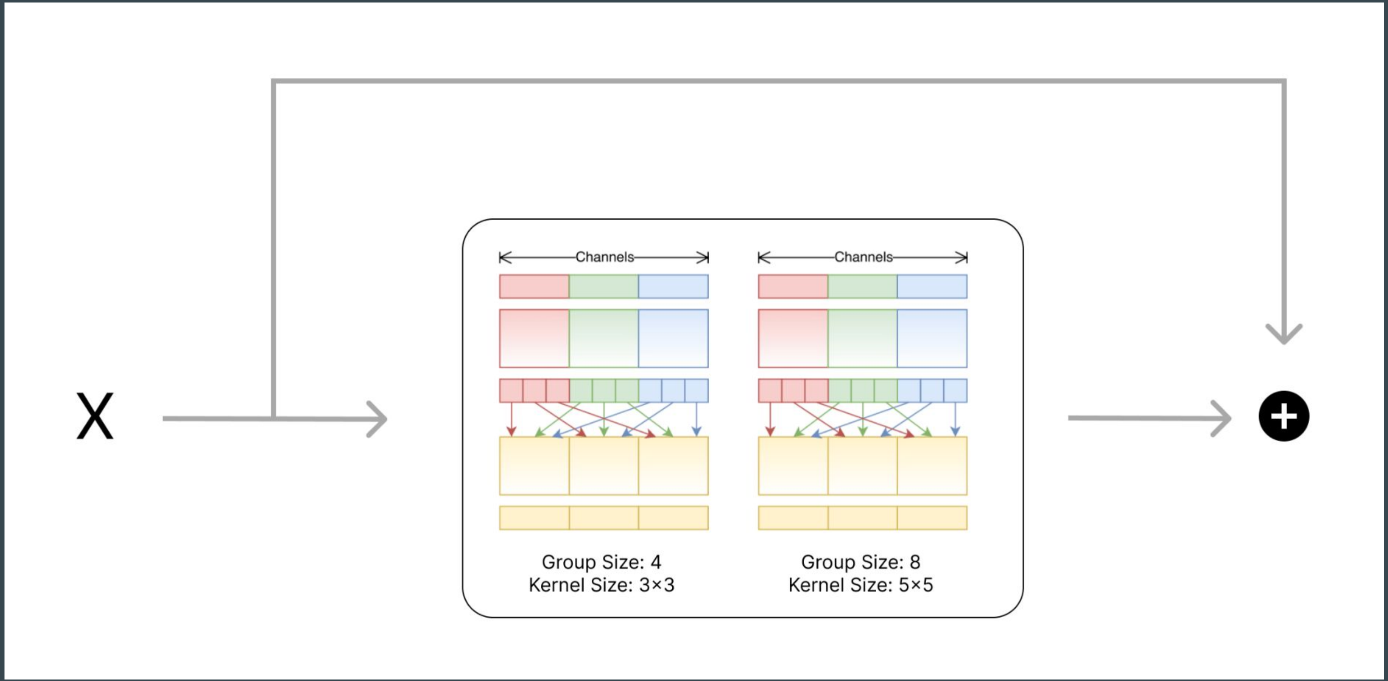
affine transformation parameters: (α, β)

Identity Preserving Model

- **Objective**: The model provides the Facial data for a face, based on that Loss component is decided. This data is the unique facial feature, which can be obtained from a Deep Convolutional Neural Network.

  We have proposed a Model Architecture for such a network, based on the implementation of XceptionNet, ShuffleNet and InceptionNet.

Model Architecture



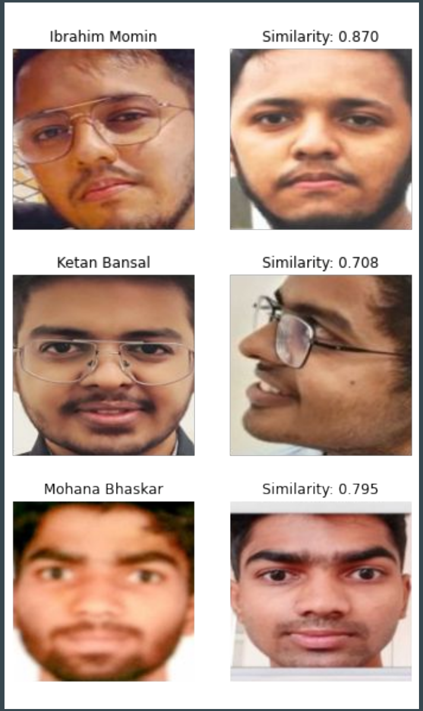
Model Architecture (contd.)

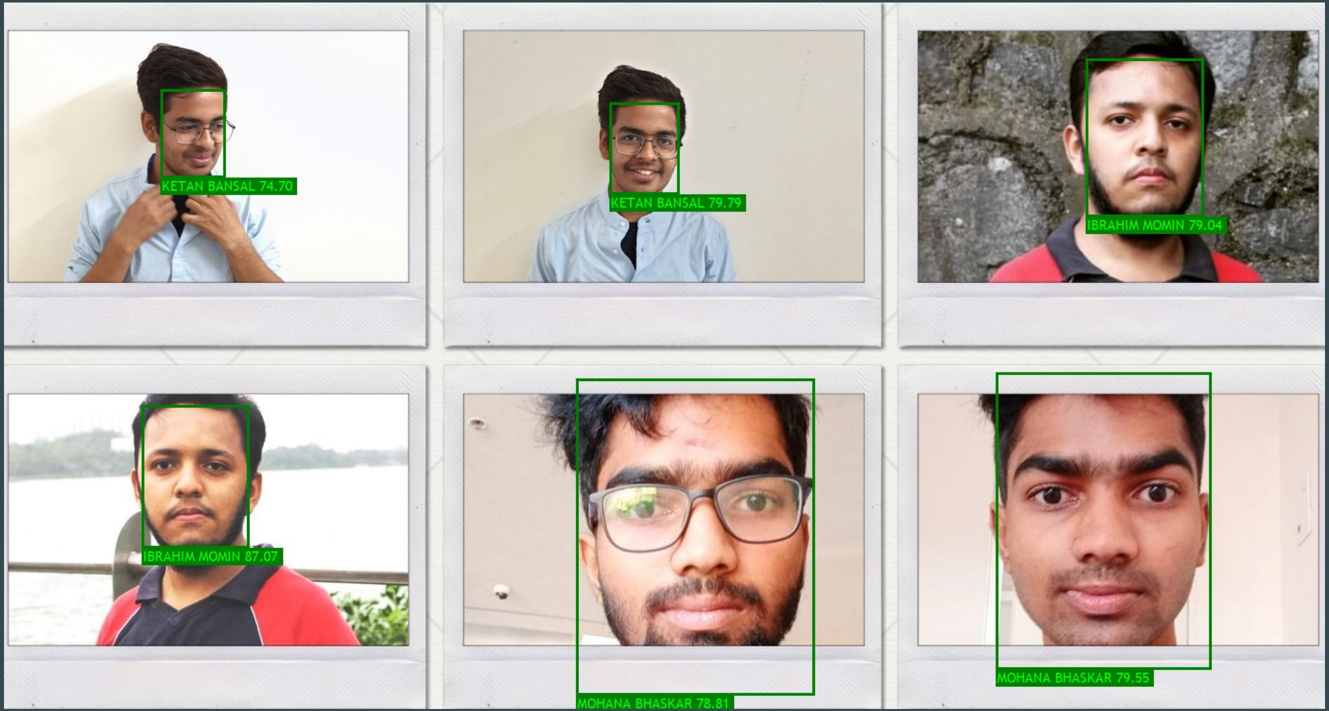
The module combines the Inception inspired modules through Shufflenet blocks, that is the grouped Depthwise convolution followed by grouped pointwise convolution from different scale of receptive fields. The single DISNet block is made of up of 2 blocks, one with 3×3 kernel size and another with 5×5 for having different scale of receptive fields, the outputs are finally added to the inputs for the residual connection.

Experimental Setup and Datasets

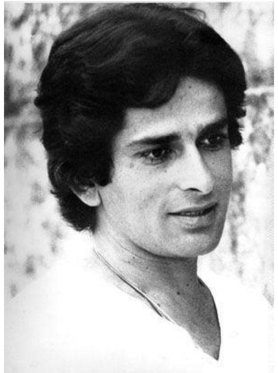
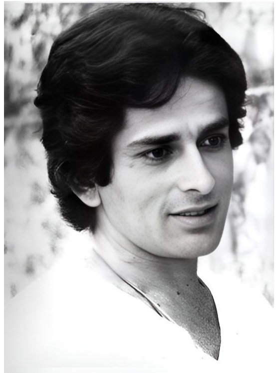
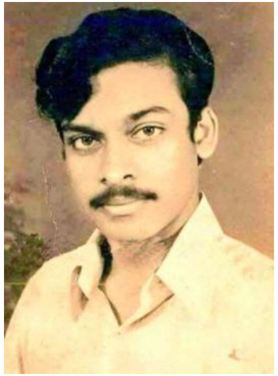
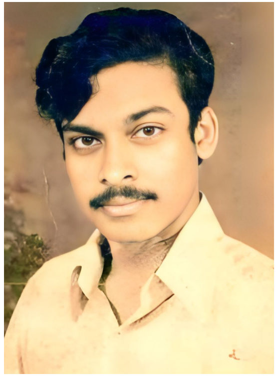
- The model training was done in 2 steps, GFP-GAN on the FFHQ dataset, which consists of 70,000 high-quality images. We resize all the images to 512 during training. GFP-GAN is trained on synthetic data that approximate to the real low-quality images and generalise to real-world images during inference.
- We adopt the pretrained StyleGAN2 with 512 outputs as our generative facial prior. The UNet for degradation removal consists of seven downsamples and seven upsamples, each with a residual block. For each CS-SFT layer, we use two convolutional layers to generate the affine parameters α and β respectively.
- Our Custom Identity Preserving Network was trained on VGGFace2 Dataset, the input size for the image is 128x128x3, the model object for training was focal loss.

Results


Ibrahim Momin
Similarity: 0.870
Ketan Bansal
Similarity: 0.708
Mohana Bhaskar
Similarity: 0.795


KETAN BANSAL 74.70
KETAN BANSAL 79.79
IBRAHIM MOMIN 79.04
IBRAHIM MOMIN 87.07
MOHANA BHASKAR 78.81
MOHANA BHASKAR 79.55

Results

Input
from *real life*

Output
Our Model

Input
from *real life*

Output
Our Model

Input
from *real life*

Output
Our Model

Results

Methods	FID	NIQE
Bicubic	148.87	10.767
ESRGAN	49.2	4.099
GFP-GAN	42.36	4.078
Ours	43.72	4.076
GT	43.43	4.292

Quantitative comparison on CelebA-Test for 4× face super-resolution

FID (Fréchet inception distance) compares the distribution of generated images with the distribution of real images that were used to train the generator.

Natural Image Quality Evaluator (NIQE) is a completely blind image quality analyzer that only makes use of measurable deviations from statistical regularities observed in natural images.

Novelty

- **Privacy Preserving:** This project focuses on using Latent Features to recognize faces while maintaining privacy.
- **Identity Preserving:** Based on ShuffleNet, InceptionNet, and Xception architecture combined the network maintains the unique identity of .
- This project can be used to obtain high quality images for Face Recognition models.

Future Directions

In future this network can be modified to work well in real-time, i.e. we can run the Face Restoration task as we receive the face in less than a second, this can help in Video surveillance like Face Recognition, Video Forensics, etc., more modifications can be on the Degradation removal module, current we are using a segmentation model instead we can use a Pixel Shuffle Network to remove the noise as used in Super Resolution.

References

[1] - GFPGAN Paper (https://arxiv.org/pdf/2101.04061.pdf) - Xintao Wang, Yu Li, Honglun Zhang, Ying Shan

[2]- StyleGAN Paper (https://drive.google.com/file/d/1fnF-QsiQeKaxF-HbvFiGtzHF_Bf3CzJu) - Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, Timo Aila

[3] - VGGFace-2 Dataset (https://www.robots.ox.ac.uk/~vgg/data/vgg_face/)

[4] - Progressive Semantic-Aware Style Transformation for Blind Face Restoration (https://arxiv.org/pdf/2009.08709.pdf)

[5] - ArcFace: Deep Face Recognition (https://arxiv.org/pdf/1801.07698v3.pdf) - Jiankang Deng, Jia Guo, Niannan Xue, Stefanos Zafeiriou

[6] - ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices (https://arxiv.org/pdf/1707.01083.pdf)

[7] - Xception: Deep Learning with Depthwise Separable Convolutions - Francois Chollet Google, Inc. (https://arxiv.org/pdf/1610.02357)

[8] - Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning (https://arxiv.org/pdf/1602.07261v2.pdf)

[9] - Siamese Neural Networks for One-shot Image Recognition(https://www.cs.cmu.edu/~rsalakhu/papers/oneshot1.pdf)

Thank You

Face Recognition