A Comparative Analysis Of Face Recognition Models On Masked Faces

Yalavarthi Bharat Chandra, Gouru Karthikeya Reddy

Abstract: Face recognition systems are among the widely used biometrics in fields such as surveillance, access controls, attendance, forensics and other security purposes. Due to current Covid-19 crisis almost everyone can be seen wearing a mask in public. This change can be very challenging for existing facial recognition systems and can make them less effective. A face mask covers significant portion of the face making facial recognition systems having less face features to recognize and on top of that face masks can also add significant noise to the image. As wearing face masks is going to be new normal it is important to understand and study how current state of art face recognition models perform in recognizing masked faces. We did a comparative analysis on four state of art deep learning models which are widely used in this field 1)VGGFace 2)FaceNet 3)OpenFace 4)DeepFace. The analysis is made on face verification task on RMFRD dataset which is largest real world masked face dataset available. We compared the models on various metrics like error rate, accuracy ,precision, verification time.

1 INTRODUCTION

Face recognition is a technology that can recognize or identify the face of a person from a digital image. It is one of the most widely studied area in biometrics because unlike other biometrics like iris, fingerprint, palm print which can be used for the identification of a human being but in all of them active cooperation of the person is required whereas in facial recognition identification can be done without the involvement of the person. It is becoming one of the most significant applications for commercials and law enforcements which include forensic identification, access controls, border surveillance [1] [2]. Over three decades of research and development in face recognition, a high performance has been achieved in many applications. Nevertheless, the recognition methods are still affected by factors such as Occlusions, illumination, facial expression, and poses. [3]Currently, face recognition is used as a technology to provide multiple security in various practices likes verification of identity, access authority, observation, to replace passwords and identity cards that are no longer safe. [3]Face recognition involves the following steps i)Face detection ii)Face alignment iii)Numerical representation iv)RecognitionFace recognition can be of two types i) Face Verification ii)Face Identification. Face Verification(One to One) is a process of verifying if two input face images belong to same individual or not. It is most used in border crossings passport control scenarios etc. Face Identification (One to many) is a process of identifying an individual by matching his face embeddings with those in the database.
Most of the recent face recognition models uses convolutional neural networks(CNN).s

1.1 CNN

There has been a significant performance improvement of models in face recognition due to usage of deep learning architectures and CNN's. In deep learning convolutional neural network(CNN) are a kind of deep neural networks, most commonly used in image recognition, object detection, image classification and facial recognition. Convolution layers will convolute the input and pass the result to the next layer. These neutrons will be able to learn the features as well as classify data. Convolution layer will extract features from a image and it prevents the relationship between the pixels by learning the features of the image using small square of the image data. Each layer in the convolution known as convolution layer contains filters known as convolution kernels. The filters is a form of matrix containing integers that are used on a subset of the input pixel values which is the same size as the kernel. The result is obtained by summing up for a single value for representing a cell like a pixel. CNN can be used for the understanding of Natural processing language and speech recognitionBecause of the current Covid-19 situation most people wear a mask in public. This can be a challenging situation for facial recognition systems because this brings significant occlusion to face and systems have far lesser features to recognize a face. It is important to understand how well the current state of art Facial recognition systems perform masked face recognition. For this reason a comparative analysis is done on existing state of art models to understand their level of performance and challenges face by them in handling this problem. We chose four most widely used state of art facial recognition models namely i)VGGFace ii)Facenet iii)OpenFace iv)DeepFace. All these models have very high accuracy on standard benchmark datasets like Labelled faces in wild(LFW) and Youtube Faces(YTF).We now test them against masked face dataset called RMFRD.

1.2 Dataset Information

For this comparative analysis we used RMFRD dataset.This is largest real world masked face dataset available for public use. A python crawler tool is used to crawl the front-face images of public figures and their corresponding masked face images from massive Internet resources. Then, they manually removed the unreasonable face images resulting from wrong correspondence. Similarly, they cropped the accurate face areas with the help of semi-automatic annotation tools, like LabelImg and LabelMe. The dataset includes 5,000 pictures of 525 people wearing masks.

2 DEEP LEARNING MODELS OVERVIEW

2.1 VGGFace

VGGFace model is developed by Visual geometry group(VGG) from Oxford university. It is one of the most popular and widely used facial recognition model. It comprises of 38 layers and is trained on 2.6M images of more than 2600 individuals. In general, the VGGFace contains thirteen convolutional layers, each layer having a special set of hybrid parameters. Each group of convolutional layers contains maxpooling layers and there are also 15 rectified linear units (ReLUs). After these layers, there are three fully connected layers namely the FC6, FC7 and FC8. The first two have 4096 channels, while FC8 which has 2622 channels are used to classify the identities. The last layer is the classifier which is a softmax layer to classify an image to which the individual face class belongs to. [4]

2.2 Facenet

FaceNet is introduced by Google researches by integrating machine learning in processing face recognition. FaceNet directly trains the face using the Euclidean space where the distance consists of similarities between facial models. The training method on FaceNet uses triplet loss that will minimize the gap of anchor and positive, also maximize the gap of anchor and negative image. FaceNet uses deep convolutional networks to optimize its embedding, compared to using intermediate bottleneck layers as a test of previous deep learning approaches. [6]FaceNet is trained using between 100M-200M training face thumbnails consisting of about 8M different identities with input size ranging from 96x96 pixels to 224x224 pixels. [7]

2.3 OpenFace

OpenFace is a open source deep learning facial recognition model developed by researchers at Carnegie Mellon University based on the Google's Facenet model. One significant advantage of openface is that it is developed with focus on real time recognition and can work seamlessly on mobile devices. So you can train a model with high accuracy with very little data. and It is implemented using Python and Torch so it can be run on CPUs or GPUs. During the training portion of the OpenFace pipeline, 500,000 images are passed through the neural net. These images are from two public datasets: CASIA-WebFace, which is comprised of 10,575 individuals for a total of 494,414 images and FaceScrub, which is made of 530 individuals with a total of 106,863 images.[9]

2.4 DeepFace

DeepFace is a deep neural network model developed by researchers at Facebook. This model is 9 layer deep and has more than 120M weights to train. These are mapped to different locally-connected layers. On contrary to the standard convolution layers, these layers do not have weight sharing deployed. This model takes a bit advanced approach than other by adding the 3D transformation and piece-wise affine transformation in the procedure, the algorithm is empowered for delivering more accurate results. [8]The model is trained with more than 4 million facial images of more than 4000 people for this purpose. This model works with 97.47 percent accuracy, which is almost equal to human eye accuracy 97.65 percent.

3 Comparative analysis

In this section four mentioned state of art models in face recognition are compared based on various metrics like accuracy, precision, error rate and verification time. Each model is tested based on their performance on face verification of masked face images from RMFRD dataset. From the 5000 images available 6384 pairs of images are chosen randomly in such a way that there are equal positive(pair of images belonging to same individual) and negative pairs(pair of images belonging to different individuals).These pairs are verified if they belong to same person or not. Face embeddings of each image is calculated and distance between them is used to classify if the pair is same person. It is important to determine a good threshold value to verify the pair of images after the distance between their embeddings are calculated. We chose the optimal threshold value based on the accuracy and precision of the model. We used the default pre trained weights of the deep learning models in this analysis. The first comparison was performed based on the error rate of models which is shown in fig[1].Error rate is calculated as (100-Accuracy).It can be inferred that VGGFace has the least error rate(31.83%) of all the models compared followed by Facenet(32.52%),DeepFace(36.22%)and OpenFace(36.81%) .Based on the error rates of the model it can concluded that VGGFace performs best.

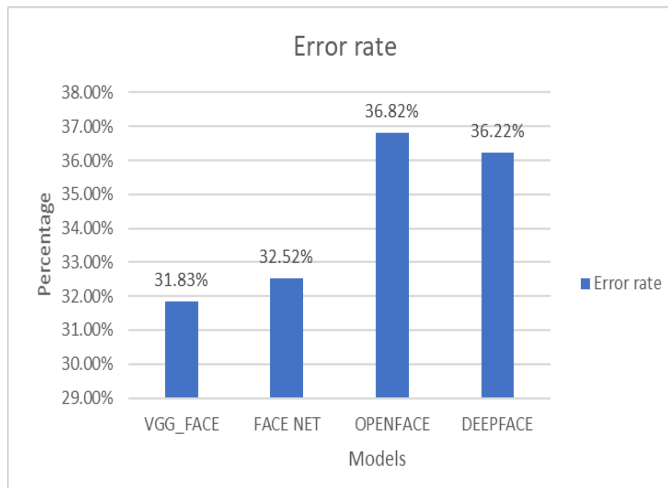

***Fig1: Error** rate of the models.*

The next comparison was made based on the accuracy of the models prediction shown in fig[2].

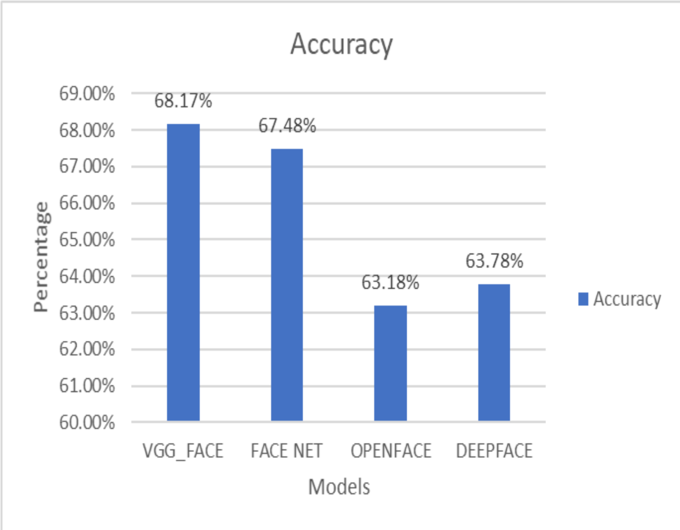

Fig 2: *Accuracy of the models.*

Next comparison was made based on the precision of the models prediction shown in fig[3]. VGGFace has the best precision of all the four models with 60.17% followed by FaceNet (59.48%), DeepFace(55.42%) and OpenFace(52.55%).

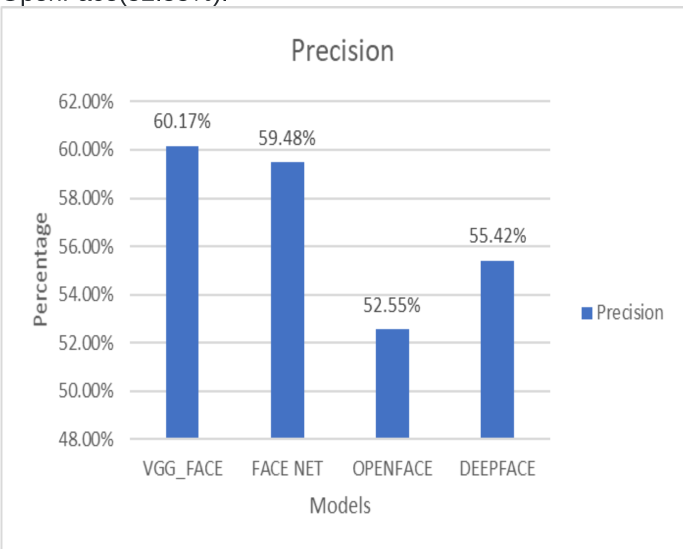

Fig 3: *Precision of the models.*

We also compared the time taken by each model to verify one pair of faces as shown in fig[4]. As seen from the bar chart VGGFace has the least time taken(0.32s) for one verification followed by OpenFace(0.34s),DeepFace(0.36),FaceNet(0.41s).

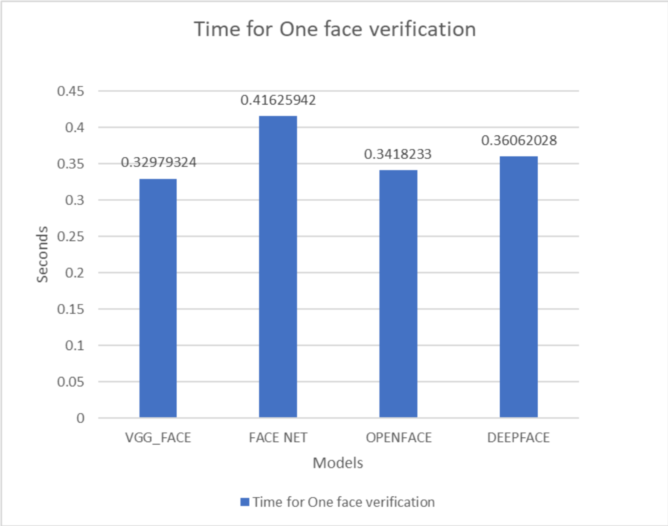

Fig 4: *Time taken by the models for one pair of face verification.*

We can infer that VGGFace performed better than other models in Error Rate, Precision and Time taken.

4 CONCLUSION

This study aims to do a comparative analysis of state of art facial recognition models with respect to their performance on masked face recognition. We compared four pre-trained models namely VGGFace, Facenet ,OpenFace and DeepFace. We used RMFRD dataset for face verification purpose. VGGFace showed the best performance in terms of accuracy(68.17), precision(60.17) and time taken(0.32s) for one verification. It can be concluded that performance of these models drops significantly while handling face images with masks. The accuracy is expected to drop further when tested in one to many face recognition systems which are used for mass surveillance. VGGFace has an accuracy of 99.13% on the benchmark LFW dataset compared to 68.17% on RMFRD dataset shows the significant difference in accuracy percentage. We can conclude that none of these pre trained models with default weights perform well enough on masked faces. The performance of these models can be further improved by methods such as transfer learning and fine tuning.

REFERENCES

[1] Mathworks. [Online]. Available: https://in.mathworks.com/ [Accessed: January 2, 2019].

[2] T. Sabharwal and R. Gupta, “Facial Marks for Enhancing Facial Recognition after Plastic Surgery,” in Proc. of the 5th International Conference on Computing for Sustainable Global Development (INDIACom-2018)”, 14th - 16th March, 2018, Delhi, India.

[3] Poses: U. Scherhag, R. Raghavendra, K.B. Raja, M. Gomez-Barrero, C. Rathgeb, C. Busch, On the vulnerability of face recognition systems towards morphed face attacks, in 2017 5th International Workshop on Biometrics and Forensics (IWBF) (April 2017), pp. 1–6VGGface

[4] Omkar, Parkhi, Andrea Vedaldi, and Andrew Zisserman. "Deep face recognition." BMVC. Vol. 1. No. 3. 2015.

[5] Deep face recognition using imperfect facial dataAuthor links open overlay panelAliElmahmudiHassanUgailMasked Face Recognition Dataset and ApplicationFacenet
[6] Face Recognition using FaceNet (Survey, Performance Test, and Comparison)
[7] FaceNet: A Unified Embedding for Face Recognition and Clustering
[8] OpenFacehttps://algorithmia.com/blog/understanding-facial-recognition-openface
[9] DeepFace https://pypi.org/project/deepface/