Webology (ISSN: 1735-188X)
Volume 18, Number 6, 2021

Face Recognition System For People With And Without Mask Using Deep Learning Model

V.Prema[1] , M.Sivajothi[2] , Grasha Jacob[3]

[#1]Research Scholar, Department of Computer Science, Rani Anna Government College for Women, Tirunelveli 627 008, Affiliation of Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli, Tamil Nadu, India 627 012.

[#2]Associate Professor, Department of Computer Science, Sri Parasakthi College for Women (Autonomous), Courtallam 627 802, Affiliation of Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli, Tamil Nadu, India 627 012.

[#3]Associate Professor and Head, Department of Computer Science, Rani Anna Government College for Women, Tirunelveli 627 008, Affiliation of Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli, Tamil Nadu, India 627 012.

ABSTRACT

Face recognition system is one of the most important biometric authentication systems frequently employed in access control and surveillance. With the superior power of three dimensional (3D) imaging sensors,3D face recognition systems have the potential to attain good recognition accuracy compared to 2D face recognition systems. However, most of the existing 3D face recognition approaches suffer from computational overhead. Recently, the COVID-19 pandemic is spreading throughout the globe which changes the lives of people and creates more problems on people’s health. Wearing a face mask in public places is an effective way to prevent viruses from spreading. This has initiated issues in current face recognition system, motivating the development of an efficient approach for recognizing both masked and non-masked faces. This paper proposes a robust 3D face recognition system based on deep learning model which is capable of recognizing both masked and non-masked faces. Convolutional neural network is designed to classify the facial images into masked and non-masked faces. Non-masked faces are recognized based on the uncovered facial features while masked faces are recognized based on the features from the preprocessed images. The recognition process is carried out using twin neural network. Performance of the proposed system is validated using four data bases such as FRGC V2.0, Texas 3D, Bosporus and Real World Masked Face Recognition Dataset (RWFRD). Experimental results showed that the proposed system provide excellent performance compared with the existing approaches.

1. INTRODUCTION

Biometric authentication system aims to identify a person based on his/her behavioral and psychological trails. Compared with other biometric authentication systems like voice, iris, palm print and fingerprint etc., Face Recognition System (FRS)establishes the identity of a person based on facial image features. It is considered as powerful authentication system for many real time applications due to its advantages over other biometrics such as non-intrusive, user friendly and social acceptance [1]. Therefore, it has become a hot topic interest for researchers and experts. Furthermore, FRS has a wide range of applications in pattern recognition, computer vision, security, access control, civilian, military, fraud detection and surveillance.

Although the past years has shown tremendous growth in 2D Face Recognition (FR), highly robust and accurate FR system in an uncontrolled environment is still an extremely challenging as the 2D facial image features is sensitive to pose, Expression and Illumination (PIE) variations. Additionally, head or body movement can cause large variations and result in missing significant features. It is demanded to prevent these problems and offer an effective solution to benefit both access control and security applications. A feasible solution is to apply three dimensional model for FR. With the advancement of 3D imaging techniques, many researchers have shifted to 3D FR due to its ability to tackle the drawbacks of 2D FR system. In addition to this, 3D images offer more geometrical data than 2D images, an 3D images are invariant to rotation scaling and illumination. Numerous methods have been developed for 3D FR. However, most of the prevailing methods are highly expensive due to the complicated preprocessing, feature extraction and matching processes [2][3].

The covid-19 pandemic is spreading throughout the globe which changes the lives of people and creates more problems on people’s health. Wearing mask in public place is a better way to avoid corona viruses from spreading. The mask has made it hard for FRS to recognize a person due to the lack of facial information. This paper intends to develop a robust and accurate 3D FR model based on deep learning network which is capable of recognizing a person based on facial features. The research also includes a study of how the FRS can recognize a person even if he/she is wearing mask. Several methods have been developed for FR using artificial intelligence and deep learning model [5][6]. Many researchers suggest to use Support Vector Machine (SVM) or Convolutional Neural Network (CNN) to achieve better results. Nevertheless, when the number of attributes is more than the number of scans, SVM does not work well. A CNN is to be better than SVM when it comes to images CNN identify features automatically, it may fail when the samples are small. To solve this issue, the proposed system adopts Twin Neural Network (TNN).The major contributions of this paper are as follows:

1. A novel preprocessing method is proposed to get 2.5D range scan from raw 3D scan for non-masked face recognition.
2. Improved deep learning model, TNN is designed to perform recognition task.
3. Rather than using 3D scan, range images are considered. These images offer more information to improve recognition accuracy.
4. Experiments are conducted on four public datasets and results proved that the proposed method reaches the best recognition rate and highest computational efficiency compared with the former methods.

The remainder of the paper is outlined as follows: Section “Review of previous approaches” provides a brief survey of latest works on FR. Section “Proposed methodology” delineates the functioning of the proposed FRS. Section “Experimental results and discussion” presents the simulation results and compares with the existing models. Section “Conclusion” summarizes the empirical findings and concludes the paper.

2. REVIEW OF PREVIOUS APPROACHES

Face recognition system has many uses in security and access control applications. In the past, 2D FRS has been utilized in such applications and has provided promising results. But, it has been shown that the recognition rate decreases when 2D FRS employed in the presence of PIE variations. These challenges are solved by 3D FRS, where 3D facial images offer more geometrical data. This section reviews methods that are developed for 3D FRS.

Recently, with the emergence of artificial intelligence and deep learning networks, FRS achieves higher Recognition Rate (RR). Kim et al. [1] developed a first 3D FR system using deep learning model, VG Gnet. Authors used 2D face model to obtain 3D representations. The direct conversion from 2D to 3D is not a good way when adopting deep learning model for 3D FR. Ratyal et al. [2] presented a 3D FRS which is based on deep learning model. Input images were preprocessed and face region was identified using CNN. SVM performed the recognition task. However, this system provided better results with high computational cost. Lei et al. [3] used local geometric signatures to identify human faces. Shi et al. [4] designed a FRS system using frenet features of geodesic curves. Recognition accuracy of the methods presented in Ref. [3] and Ref. [4] depends on the feature extraction process. The main drawback is that these methods require efficient algorithm for feature extraction in order to achieve better RR. A deep learning approach for 3D FR was designed by Cai et al. [7].In this approach, 3D facial images were preprocessed to obtain 2.5D range images and then nose tip was extracted for alignment. A deep learning model, Res Net was designed to recognize the images. This approach provided better results than that of the other machine learning models. Vu et al. [8] presented a masked face recognition system using neural network. Facial images were preprocessed and local binary patterns were extracted. The computed patterns were fed as input to the neural network model to recognize the person. Saxena et al. [9] developed a masked FRS using Siamese neural network. This method achieved good recognition accuracy. Li et al. [11] proposed a face recognition approach based on Siamese neural network and achieved better recognition accuracy than that of the other models.

3. PROPOSED METHODOLOGY

The overall flowchart of the proposed FRS is evinced in Figure.1. As shown in Figure.1, the proposed system first detects whether the person is wearing a mask or not. If the person is wearing a mask, the image is sent to Masked FRS (MFRS) or to FRS. The proposed system uses standard CNN and TNN to achieve high recognition rate. Standard CNN is used for detecting if the person’s is wearing a face mask or not whereas TNN is employed for recognition. The overview of the proposed FRS is depicted in Figure.2.

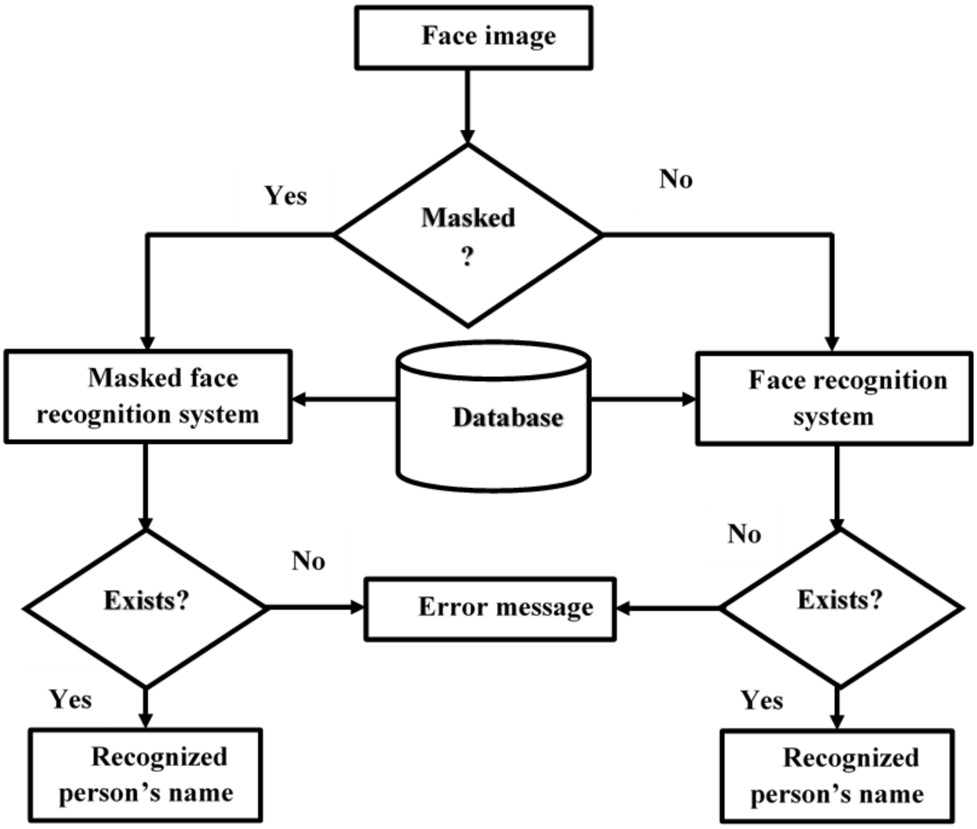

Figure.1. Flowchart of the proposed system

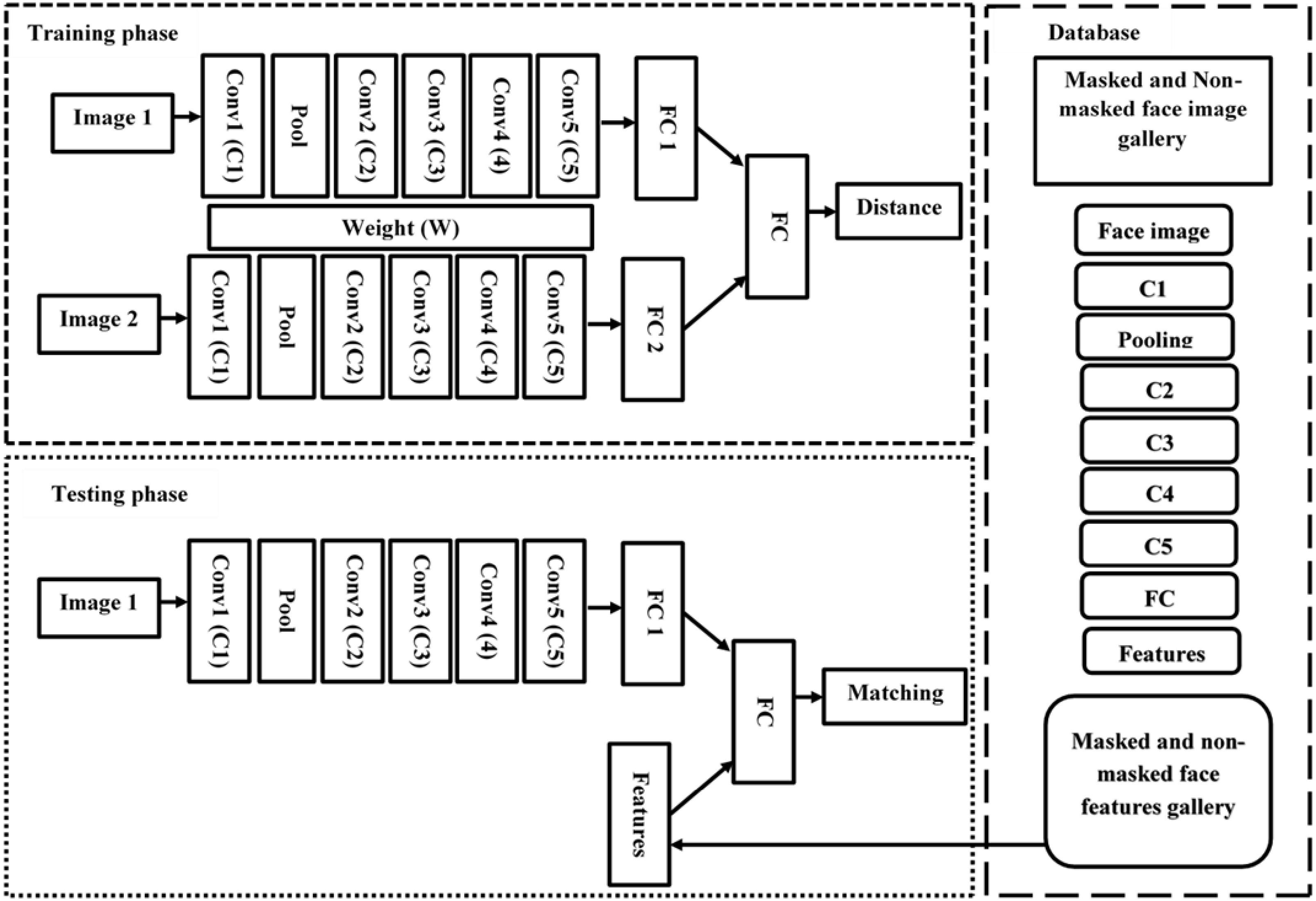

Figure.2 Proposed face recognition system

3.1 Face mask detection

The CNN architecture has been successfully used in pattern classification and computer vision fields due to its characteristics of capturing spatial and temporal dependencies in an image via the application of filters. In this work, CNN is employed for masked face detection. The CNN is built with various layers such as convolution layer, pooling layer, flatten (F), dropout (DO) and dense (D) layer, as illustrated in Figure.3. Convolutional layer applies a convolution operation to the input image and passing the result to subsampling layer. The subsampling layer is responsible for reducing the size of convolved features. Flatten layer converts the data into 1D array for inputting it to the dropout layer. The dropout is used to prevent overfitting problem. Dense layer uses soft max function to output a vector that provides probability of each of the masked and non-masked faces. The network is trained with Adam optimizer. Based on which probability is higher, the recognition system will be chosen.

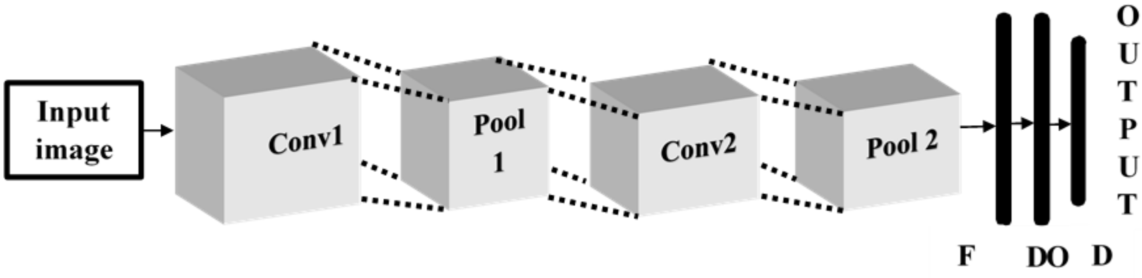

Figure.3 Structure of Convolutional neural network for masked and non-masked face classification task

3.2 Face recognition system

If the person is not wearing a face mask, then the image is sent to the FRS which uses TNN for recognizing the person. TNN works efficiently in such a case as it differentiates two classes by computing similarity score instead of detecting features of a class. It can show whether a pair of images belong to the same person or not by analyzing pairs of images. In this work, a preprocessing method is proposed to obtain Region of interest from the raw image. Five landmarks namely nose tip, center of pupils and corners of mouth are detected using Dlib [10]. Five landmarks on the 3D image is calculated based on the correspondence between the 2D and 3D scan. A sphere on the nose tip with a radius of 100mm is drawn to obtain the face region and then Principal Component Analysis (PCA) pose correction algorithm is applied to normalize the pose. A nose tip refine process is done using depth information to control the ambiguity of nose-tip detection. A sphere centered on the nose tip with a radius of 30mm is drawn to cut nose region and selected points whose distance is less than a threshold of 20mm. The cropped image is filtered by hybrid median filter and interpolated. Finally, a new nose tip (maximum point) is selected and normalized face mesh onto a square grid of 320x320 mm with a uniform resolution of 1mm.Preprocessing steps of the proposed FRS is shown in Figure.4.

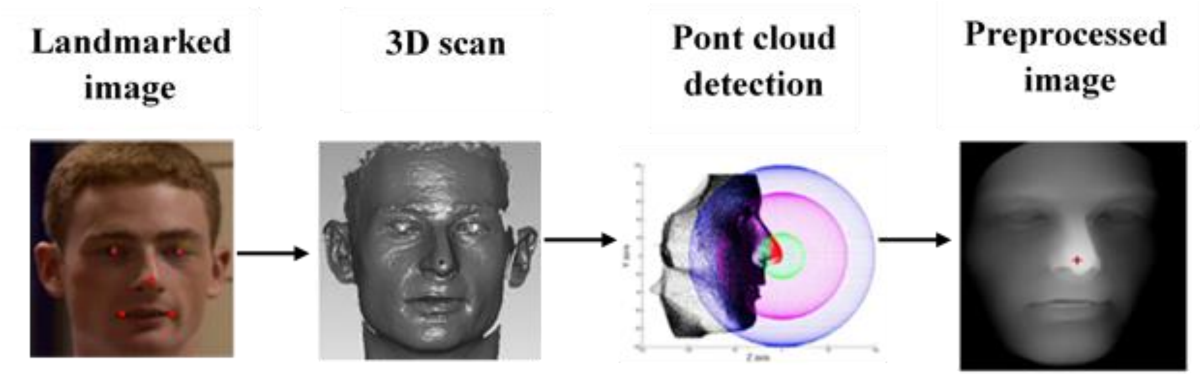

Figure.4 Preprocessing

After preprocessing, the proposed FRS is implemented by inputting a pair of scans: the preprocessed image and an image from the database to the TNN. The TNN process the images and compare the features. Distance between two features is computed and the obtained distance represents the similarity score between a pair of images. The similarity score of same person is low whereas the similarity score of different person is high. Keeping this score as a base, the proposed system compares the input images with other images in the database. Finally, the proposed system provides a recognized person name based on the features which have low similarity score.

3.3 Masked face recognition system

If the person is wearing a facial mask conventional FRS would not work and would provide low recognition rate. In this case, features on the upper part such as eyes and eyebrows are selected as features to recognize the person. The developed MFRS will be on two aspects. One database possesses both masked and unmasked faces while other one has uncovered facial features. The proposed system compares the test image features with the database. If the image is identified, then it displays a person's name. Else, if the mage features are not registered in the database, it will show an error.

3.4 Twin Neural Network

Many research studies have showed that standard CNN could effectively obtain the salient features to achieve high classification accuracy. However, as the number of classes increases, the classification accuracy decreases and it is difficult to construct an output layer for all classes. Therefore, TNN is used to prevent this issue caused by the small number of images. There are two major differences between TNN and CNN. With respect to network architecture, TNN builds two convolutional channels in the form of weight sharing and the two channels are employed to compute the features from two different scans. In terms of objective function, the TNN replaces the entropy function widely used in classification task via a distance measurement method, which increases the network's interpretability. Structure of TNN used in this system is depicted in Figure. 5. As shown in Figure.5, there are two channels with shared weights. The shared weight represented by W which are utilized to extract relevant facial features. The parameters along with their size of the TNN are tabulated in Table.1.

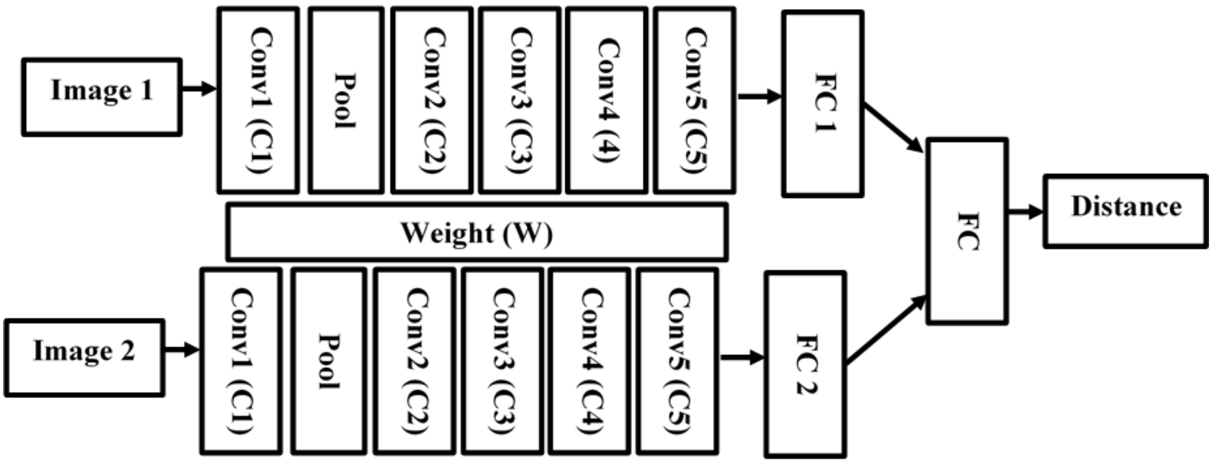

Figure. 5 Structure of TNN

Table.1 Parameters of the TNN

Layer name	Size	Channel 1	Channel 2
C1	112 X 112X 64	7 X 7, 64, stride 2 3 X 3 Max pool, stride 2	7 X 7,64, stride 2 3 X 3 Max pool, stride 2
C2	56 X 56 X 64	$\begin{bmatrix} 3 X 3, & 64 \\ 3 X 3, & 64 \end{bmatrix}$ X 2	$\begin{bmatrix} 3 X 3, & 64 \\ 3 X 3, & 64 \end{bmatrix}$ X 2
C3	28 X 28 X 128	$\begin{bmatrix} 3 X 3, & 128 \\ 3 X 3, & 128 \end{bmatrix}$ X 2	$\begin{bmatrix} 3 X 3, & 128 \\ 3 X 3, & 128 \end{bmatrix}$ X 2
C4	14 X 14 X 256	$\begin{bmatrix} 3 X 3, & 256 \\ 3 X 3, & 256 \end{bmatrix}$ X 2	$\begin{bmatrix} 3 X 3, & 256 \\ 3 X 3, & 256 \end{bmatrix}$ X 2
C5	7 X 7 X 512	$\begin{bmatrix} 3 X 3, & 512 \\ 3 X 3, & 512 \end{bmatrix}$ X 2	$\begin{bmatrix} 3 X 3, & 512 \\ 3 X 3, & 512 \end{bmatrix}$ X 2
FC1	500 X 1	Avg. pool 500-d fc	Avg. pool 500-d fc
FC2	1 X 1	Avg. pool 1-d fc	

4. EXPERIMENTAL RESULTS AND DISCUSSION

This section discusses the experimental results of the proposed face recognition systems. It also provides details of the dataset used for validation, performance measuring metrics utilized of evaluation and compares the recognition performance with the prevailing methods to prove the effectiveness of the proposed method more intuitively.

4.1 Simulation environment

The proposed face recognition system has been implemented on MATLAB2019a platform and executed in Intel core i5 processor with 2.5GHz speed with 16GB RAM and validated using publicly available dataset.

4.2 Description of dataset

The developed FRS has been validated using four datasets namely. The facial datasets are (i)FRGC V2.0 dataset (ii)Texas 3D dataset (iii) Bosphoros dataset and (iv) Real world Masked Face Recognition Dataset (RMFRD).

FRGC V2.0 is the one the popular and widely used databases for 3D FR. According to the image acquisition period, this database is grouped into three sets: (i) spring 2003 set has 943 scans of 277 individuals (ii) fall 2003 set and (iii) spring 2004 set has 4007 scans of 466 individuals. Spring2003 is adopted for training and remaining two groups for testing. FRGC V2.0 data has variations in both illumination and expressions and limited pose variations. The data base has both male and female facial images with age 18 years and above. Example images from the FRGC V2.0 data are shown in Figure. 6.

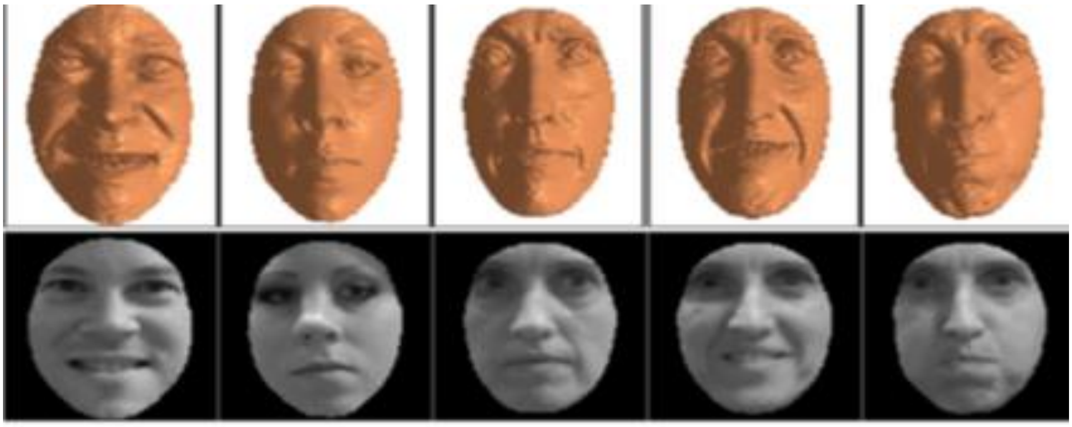

Figure. 6 Sample images of the FRGC V2.0 face database

The Texas 3D face database consists of 1149 pairs of color and range images of 118 adult human subjects. These images are acquired employing a stereo imaging system manufactured by 3Q Technologies (Atlanta, GA) at a very high spatial resolution of 0.32 mm along the x, y, and z dimensions. The database includes facial images of both male and female, aged between 25 and 75. The facial expressions present are smiling or talking faces with open/closed mouths and/or closed eyes. The neutral faces are emotionless. Example face images from the Texas3D database are depicted in Figure.7.

The Bosphoros data base composed of 4666 images acquired from 105 subjects among which 60 males and 45 females, aged between 25 and 35. These image have been captured under expression variations, pose changes and occlusions. The 2902 scans contain expression variations. Sample images from the Bosphoros database are shown in Figure.8.

Figure.7 Sample images of the Texas 3D database

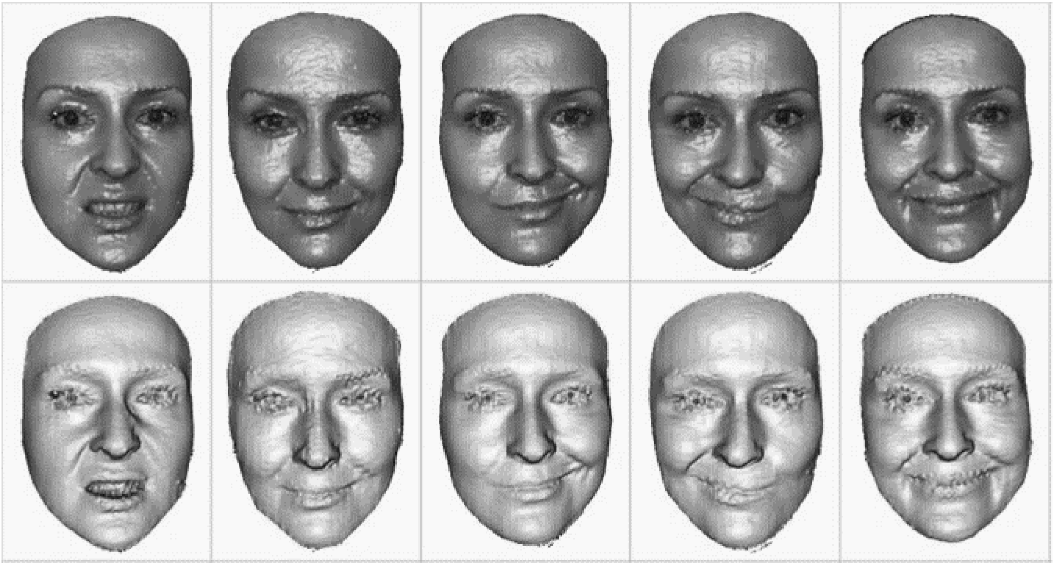

Figure.8 Sample images of the Bosphoros database

Figure.9 Sample images of the RMFRD database

Real world masked face database is a collection of masked facial image created to enhance the recognition rate of the available FRS on the masked faces during the covid-19 pandemic. In this investigation, Real world Masked Face recognition Dataset (RMFRD) is used to validate the recognition performance. The RMFRD consists of 5000 facial images of 525 subjects with masks and 90,000 facial images without masks which represents 525 subjects. Figure. 9 shows some sample pairs of face images from RMFRD database.

4.3.1 Performance analysis and comparison of the proposed FRS with the existing methods on the FRGC V2.0 data

In this section, numerical results obtained by the proposed FRS was presented. In the FRGC V2.0 facial database, the neutral image of each subject, 466 scans formed a gallery set and 3 probe sets are formed namely (i) Neutral (N) Vs. Neutral (N Vs. N) (1944) (ii) N VS. Non-Neutral (N Vs. NN) (1597) and (iii) N Vs. all (N Vs. all) (3541). To demonstrate the efficiency of the proposed FRS, simulations were conducted on the FRGC V2.0 data and results were tabulated in Table.2. Recognition performance of the proposed system compared with the existing methods to prove its superior power, given in Table.1. From the Table.2, it is noted that the proposed system provided better results than that of the existing methods by achieving 100% Verification Rate (VR) at 0.1% False Acceptance Rate (FAR) in NN experiment, 99.5 % in N Vs. NN and 99.9% in N Vs. all experiment. Most of the existing methods have used CNN for developing FRS. But, these methods failed to achieve good results. The proposed FRS used TNN which helps to reach higher VR compared with the exiting methods. The ROC and CMC curves are represented in the form of graph in Figure .10 and Figure.11 respectively.

Table.2 Performance comparison between the proposed system and the existing methods on FRGC V2.0 data at 0.1% FAR

Researchers	N vs. N (%)	N vs. NN (%)	N vs. all (%)
Ratyal et al. [2]	**100**	**96.7**	**98.4**
Cai et al. [7]	99.2	97.8	98.9
Li et al.[11]	99	94.3	96.7
Proposed	**100**	**99.5**	**99.9**

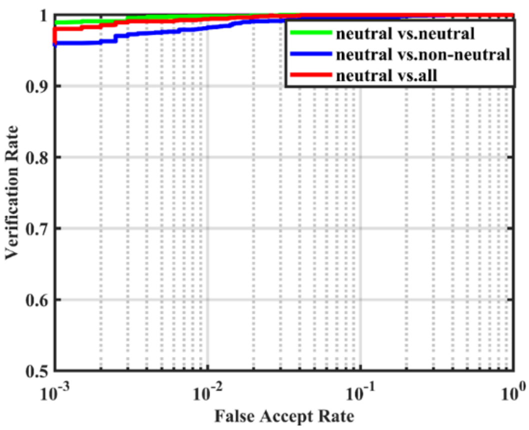

Figure.10 ROC curves of N vs. N, N vs. NN and N vs.all experiments on the FRGC V2.0

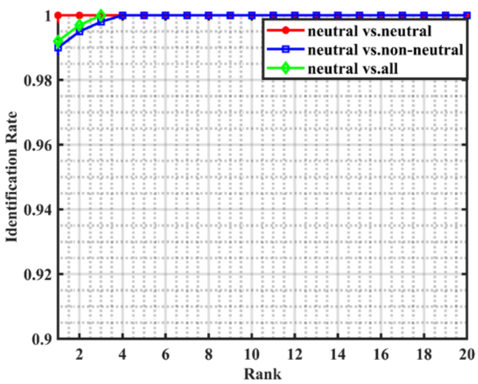

Figure.11 CMC curves of N vs. N, N vs. NN and N vs.all experiments on the FRGC V2.0

4.3.2 Performance analysis and comparison of the proposed FRS with the existing methods on the Texas 3D data

In the Texas 3D facial images, the neutral images of each individual,105 facial images form a gallery set and 3 probe sets are formed such as (i)N Vs. N (480) (ii)N Vs. NN (183) and (iii) N Vs. all (663). The proposed FRS trained using and validated with probe sets and the recognition performance reported in Table.3. Further to this, efficiency of the proposed FRS compared with the existing approaches and tabulated in Table.2 in terms of VR at 0.1% FAR. The proposed FRS attained VR of 100% for N Vs. N, 99.9% for N Vs. NN and 100% for N Vs. all case. It is noted from the Table.3 that the performance of the proposed FRS supersedes other methods in terms of VR for N Vs. all experiment. This is because of the use of the twin neural network. Twin neural network uses two channels to extract features from two images and compares the similarity after training which makes it superior that other methods. The ROC and CMC curves are drawn in Figure .12 and Figure.13 respectively.

Table.3 Performance comparison between the proposed system and the existing methods on Texas 3Ddata at 0.1% FAR

Researchers	N vs. all (%)
Gupta et al. [12]	95.80
Lv et al.[13]	93
Shi et al. [14]	96.83
Proposed	100

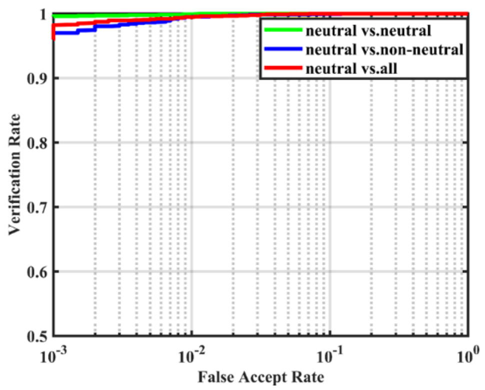

Figure.12 ROC curves of N vs. N, N vs. NN and N vs.all experiments on the Texas 3D data

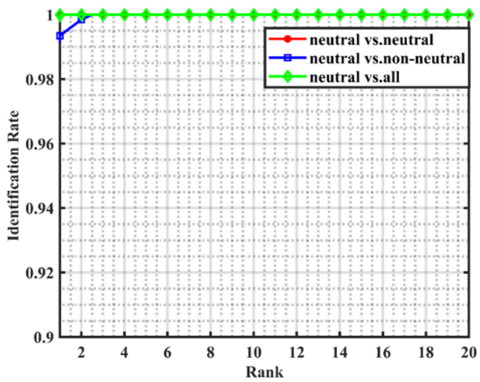

Figure.13 CMC curves of N vs. N, N vs. NN and N vs.all experiments on the Texas 3D data

4.3.3 Performance analysis and comparison of the proposed FRS with the existing methods on the Bosphoros database

Neutral scan of each individual,105 scans form a gallery set and 3 probe sets are formed such as (i)N Vs. N (194) (ii) N Vs. NN (2603) and (iii) N Vs. all (2797). To demonstrate the effectiveness, the proposed system was validated using the Bosphoros database and obtained results were presented in Table.4. Performance of the proposed system also compared with the existing approaches with respect to VR at 0.1% FAR, presented in Table.4. From the Table.4, it is found that the proposed FRS reached 100% VR at 0.1% FAR for N Vs. N,99.9 % for N Vs. NN and 100% for N Vs. all, which are higher than the other FRS. Figure.14 and Figure 15 depicts the ROC and CMC curves on the Bosphoros database for all three cases respectively. Based on the comparison, it is proved that the proposed FRS provided promising results compared to the existing models considered for comparison by achieving higher VR for all the cases considered.

Table.4 Performance comparison between the proposed system and the earlier methods on Bosphoros data at 0.1% FAR

Researchers	N vs. N (%)	N vs. NN (%)	N vs. all (%)
Ratyal et al.[1]	98	94.5	96
Cai et al.(7)	100	98.30	98.39
Li et al.[11]	98.1	96.3	95.4
Proposed	100	99.9	100

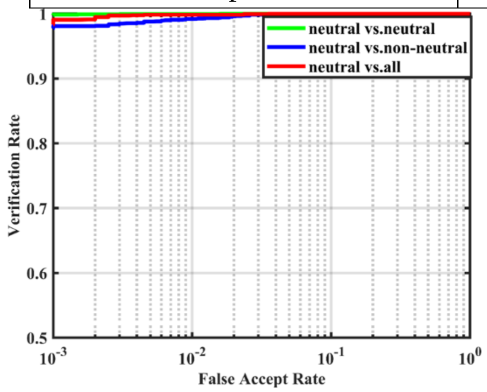

Figure.14 ROC curves of N vs. N, N vs. NN and N vs.all experiments on the Bosphoros data

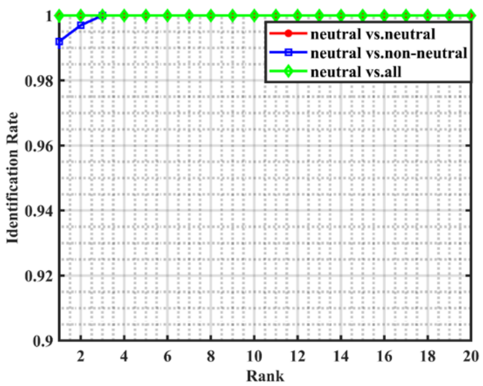

Figure.15 CMC curves of N vs. N, N vs. NN and N vs.all experiments on the Bosphoros data

4.3.4 Performance analysis and comparison of the proposed FRS with the existing methods on the RMFRD data

Table.5 presents the performance comparison of the proposed FRS with the existing methods with respect to VR at 0.1% FAR. From the Table.5, it is observed that the proposed FRS posed the VR of 99% which is higher than other methods.

Table.5 Performance comparison between the proposed system and the earlier methods on RMFRD data at 0.1% FAR

Researchers	RR (%)
Vu et al.[8]	88
Saxena et al.(9)	91
Hariri et al.[15]	93
Proposed	99

5. CONCLUSION

In this article, a novel system based on twin neural network with applications with 3D FR is proposed and validated. The face recognition system was implemented for both masked face and non-masked face recognition. Primarily, the CNN was implemented to classify the input images into masked face and non-masked face. If the person is wearing a mask, the image was passed to masked face recognition system. Else the image was sent to face recognition system. Masked face recognition system recognized faces based on the uncovered feature while face recognition system recognized the faces based on similarity score. Several experiments were conducted using four publicly available databases namely FRGC V2.0, Texas 3D, Bosporus and RMFRD. Empirical findings proved that the proposed face recognition system is capable of recognizing the facial images with higher accuracy. Additionally, experimental results proved that the twin neural network significantly improved the accuracy. As a future research direction, metaheuristic algorithm will be explored to further improve accuracy.

REFERENCES

[1]. D.Kim, M.Hernandez, J.Choi and G.Medioni (2017) Deep 3D face identification. IEEE Int. Joint Conf. on Biomets.,133-142.

[2]. N. Ratyal, I. A. Taj, M. Sajid and A. Mahmood (2019), Deeply learned pose invariant image analysis with applications in 3D face recognition, Math. Prob. in Engg., 2019,1-21.
[3]. Y. Lei, M. Bennamoun, M. Hayat and Y. Guo (2014), An efficient 3D face recognition approach using local geometrical signatures, Patt. Recogn.,509-524.
[4]. B. Shi, H. Zhang, R. Zhang and S. Zhan (2019), An efficient 3D face recognition approach using frenet feature of isodesic curves J. Vis. Comm. Image R. 59, 455-460.
[5]. P.J. Philips, P.J.Flynn and T.Scruggs et al.(2005),Overview of the face recognition grand challenge. In Proc. Of the 2005 IEEE Comp. Society Conf. on Comp. Visi. And Patt. Recogn., CVPR 2005, 947-954.
[6]. A.Savran, N.Alyuz and H.Dibeklioglu et al. (2008), Bosphoros database for 3D face analysis. In Biomet. And Identity Manag., Vol.5372 of Lecture Notes on Comp. Sci.,47-56, Springer Berlin Heidelberg, Germany.
[7]. Y. Cai, Y. Lei, M. Yang, Z. You and S. Shan (2019) A fast and robust 3D face recognition approach based on deeply learned face recognition', Neurocomput., 363 (1), 375-397.
[8]. H.N.Vu, M.H.Nguyen and C.Pham (2021) Masked face recognition with convolutional neural networks and local binary patterns, Applied Intelligence,1-14.
[9]. N.Saxena and D.Varshney (2021) Smart home security solutions using facial authentication and speaker recognition through artificial neural networks, Int. J. of Cog. Compu. in Engg.,2, 154-164.
[10]. D. E. King, Dlib-m/: A machine learning toolkit, JMLR. Org, 2009.
[11]. Z.Li, H.Zou, X.Sun, T.Zhu and C.Ni (2021) 3D expression invariant face verification based on transfer learning and Siamese network for small sample size,Electronics,10,1-14.
[12]. K.R. Gupta, M.K. Castleman, A.C. Markey and A. C. Bovik (2010),'Texas 3D Face Recognition Database, IEEE Southwest Symposium on Image Analysis and Interpretation, 97-100.
[13]. C.Lv and J. Zhao (2018) 3D Face Recognition based on Local Conformal Parameterization and Iso-Geodesic Stripes Analysis, Mathematical Problems in Engineering,1-10.
[14]. L.Shi , X.Wang and Y.Shen (2020) Research on 3D Face Recognition Method Based on LBP and SVM, Optik,.
[15]. W.Hariri (2021)Efficient masked face recognition during the COVID-19 pandemic. arXiv:2105.03026v1 [cs.CV] 7 May 2021