International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Transition from Holistic to Deep learning Face Recognition Methods

R. S. Sabeenian, J. Harirajkumar, Lizzie D' Cruz

Abstract*: *Face recognition, the fastest growing biometric technology of computer vision, made a breakthrough in the field of security, healthcare, access control and marketing etc. This technology helps in automatically discern and identify the faces for authentication by comparing available digital image of faces. Various algorithms have been developed for enhancing the performance of face recognition system. The face authentication system entails three major steps, face detection, feature extraction and face recognition. This paper provides some of the major milestones of face representation for recognition like holistic learning approach, feature based approach, hybrid approach and deep learning approach. The various techniques under these categories are reviewed. Finally, implemented face recognition using convolution neural network (CNN). In this method, the image is captured through webcam for the dataset preparation. The detection is carried out by CNN cascade, followed by face landmark and face embedding by FaceNet CNN. Recognition of face is performed after training the network. Implemented faces recognition successfully and accurately for smaller dataset.

***Keywords*: *face recognition, holistic learning, feature extraction, deep learning, convolution Neural Network*.**

I. INTRODUCTION

Face recognition (FR) is used to ascertain and verify the presence of face of human. The process are involved in the face detection are verification and recognition of face. In the face verification, mapping of the given face is done one to one against the known identity. In the second mode, face identification is done by mapping of one face to many faces available in the dataset. Various method and algorithms are developed which has improved the ability of accurately recognizing the face. The FR system consists of the following major steps:

1. Creation of dataset of Image.
2. Face detection
3. Extraction of features
4. Recognition

Manuscript published on November 30, 2019.
* Correspondence Author
Dr.R.S.Sabeenian*, Professor and Head in ECE Department in Sona College of Technology, Salem, Tamil Nadu, India.
J.Harirajkumar, Associate Professor, Department of ECE, Sona College of Technology, Salem
Lizzie D' Cruz , Lecturer (SG) in Dr. B. R. Ambedkar Institute of technology

Creation of dataset – The dataset is prepared by capturing the image through webcam, images or pre-trained data sets available in net.
Face detection – The images in the data set is compared, faces are located automatically from the given image or video and bounding box is made around the face.
Extraction of Features – After detection of face, feature extraction is performed by dimension reduction, information packing and noise cleaning etc. If face patches are not extracted then recognition system will be too large and robust to build.
Recognition - The faces are labeled by giving name for the detected face. Facial features are mapped from the image or video and which is compared with the available images in the database. When the matching is confirmed, the face will be identified.

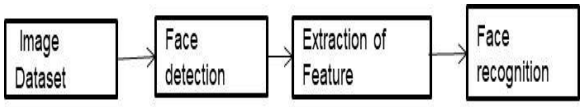


Fig.1: Components of Face Recognition System

The different methods of face recognition approaches are: holistic learning approach, feature based approach, hybrid approach [1] and deep learning approach.
Face recognition has wide application area like access control, law enforcement, security, identification of criminal, unlocking of phone, finding lost person, aid for helping the blind, forensic investigation aid, persons identification through social media platforms, diagnoses of diseases, schools and college attendance authentication system, secure payment transaction, boarding pass generation in airports and many others.
In the second section of this paper, literature review has been carried out on different approaches of face recognition. In the third section, implementation has been done for face recognition using CNN method.

II. LITERATURE REVIEW ON DIFFERENT APPROACHES OF FACE RECOGNITION

A. Holistic Approach

In this approach, the entire face area is considered as a single feature which is further matched with available face image in dataset. The Holistic approach can be either statistical approach or Artificial Intelligence approach. In the first approach the density of the face image is calculated and the obtained set value of density is compared with the values of the dataset images. In the second approach tools like neural network is used [1].

Some of the holistic learning approach are Eigen face method, Principle component analysis (PCA)method, Fisher face approach, Local binary pattern(LBP) approach, Independent component analysis, Linear discriminator analysis.

Face recognition, performed in [2] by Eigen face, fisher face as well as local binary pattern with analysis of success rate. In the proposed system viola John face detection method is used. Depicted better performance and speed with LBP approach. The filtering system used for all three methods, based on Euclidean distance calculation. Varying result of success rate has been proposed due to pose, light and expression in the case of Eigen face and fisher face method. Another model illustrated in [4] uses Eigen face method with different condition and 79% and 65% accuracy obtained. Comparative study made on ICA sensitivity to space dimension and discriminant performance of ICA is proposed in [3].

A system with fisher face approach, wherein the face space is reduced by PCA and fisher's linear discriminator has been used for feature extraction. The recognition of face is performed by Euclidean space. The processing is carried out Matlab programming language. Adobe Photoshop Cs4 program is used for preprocessing of image. The results claims, 93% of accuracy on 73 test image with the indication of the reasons of failure as scaling factor and pose [5].

Face recognition by linear binary pattern method Illustrated in [6] portrayed a system in which face image has been divided into 7x7 equal size cell. A decimal value has been assigned to the pixels. The threshold value is decided based on the weights of the neighboring cell. When the matching of face is found same as per the dataset, the door of the classroom will open as the door is connected with the servo motor and the attendance is entered in the MySQL database. 95% accuracy obtained on 11 person's image.

Fig2: (a) Dividing the image to 7x7 equally size cell, (b) weighting schemes after encoding the histogram for each of the cells [6]

A FR system provided in [7], uses Viola and John's algorithm for face detection, which is followed by FR by both Eigen face and PCA method to overcome the issues, related to light and pose.

B. Feature Based Approach

In this approach the features like, mouth, nose, chin are extracted and fed to the classifier. Some of the feature based approaches are Elastic Bunch Graph Matching (EBGM), template matching method. In template matching, patterns are represented by predefined sample, model, texture, pixel etc. correlation or distance is used as recognition function to locate the given identity. In template method the face is divided into eye, mouth and face contour and mostly used detection method is edge detection. FR based on geometrical features with Euclidean distance proposed by *Kanade* [8] involves computation of geometrical features such as nose width and length of the mouth position, shape of chin etc. from the given face. He has matched the features with the features of known individual. Closest match was achieved by Euclidean distance. The test was performed using 800 photographs. Experiments were carried with 20 persons.

In the EBGM approach, the faces are epitomized as graphs with nodes situated at fiducial points. Labeling has been done at edges with 2D distance vectors. Jet Gabor coefficients from each points were extracted manually by locating fiducial points like eyes, nose etc. on the image. Dynamic link architecture is created by Gabor wavelet transform which shows the face in elastic grid. The result of Gabor filter is provided for the feature extraction and shape detection [9].

C. Hybrid Approach

Both the holistic learning approach and feature extraction approach together is used in hybrid approach. In an example of hybrid method, depicts two modes. The first mode is the training mode and the second mode is classification. In the training of image, training set, feature extraction by principal component analysis (PCA) and independent component analysis (ICA) is proposed. Back propagation neural network (BPNN) is used for training and is done in parallel to obtain different face classes by partitioning the feature space. Classification is further proposed by PCA-BPNN or ICA-BPNN based on the face class. Testing is performed on ORL database. The result proposed reflects PCA implementation weaker than ICA [10].

Another example proposed in [11] proposes PCA and Artificial Neural Network. Recognition is performed based on the light, emotions and noises. The recognition rate observed for ideal case was 93.5% and for noise affected image, the recognition rate observed was 85%. Satisfactory performance has been proposed using feed forward algorithm. Similarly, hybrid approach has been used for detection and recognition by PCA and K-mean clustering in [12].

Another promising example of hybrid model, which combines the properties of feature and holistic based method along with Markov random field (MRF). In this work image of the face is divided into segments, and the model shows a relationship between segments or patch ID and face patches. ORL database, Yale database and FERET database for four methods, Markov random field (MRF), PCA, linear Discriminator Analysis (LDA) and Neural Network (NN). The proposed example shows evidence of having better performance than PCA and LDA methods [13].

D. Deep Learning Approach

Deep learning approach has made a tremendous impact on research based on FR. It offers better performance with respect to speed and accuracy. Most common method of deep learning method is convolution neural network(CNN)[14], a feed forward neural network constituted with many layers such as convolution layer, pooling, rectified linear unit(Re LU layer) and fully connected layers. Convolution layer has independent filters and filters are convolved with the image, are responsible for extraction of features. The CNN layer is followed by pooling layers, responsible for reducing the dimension, which in turn reduces number of parameter and computation in the network.

The process of pooling shortens the training time and controls over fitting. One of the commonly used pooling methods is max pooling. Next layer is rectified linear unit, an activation function with not vanishing gradient, computationally efficient, gives better convergence performance used in almost all deep learning approach.
The output of convolution and pooling is received by fully connected layer to predict the best label to describe the face image.

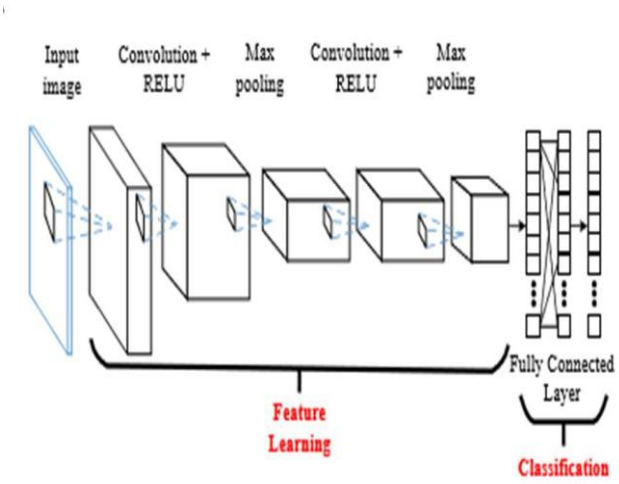


Fig 3: Convolution Neural Network [14]

Some of the models used to implementing deep learning based face recognition are: DeepFace, DeepID, FaceNet, AlexNet based CNN method, Improved AlexNet CNN method.
In the vGGface2 based approach, large face dataset has been introduced with the objective of having larger number of identity with more number of images for each identity, label noise minimization with the coverage of pose, age and ethnicity. Training of face is implemented by REsNet-50. Comparison between the VGG face2 and MSceleb-1M dataset based recognition realized shows improved performance for VGGface2. Research demonstrated and proposed performance improvement with IJB-A and IJB-B face recognition benchmarks [15].
Face similarity is measured by using FaceNet model, provides 99.63% accuracy which is captured using LFW dataset and 95.12% achieved using YouTube Faces DB.
Deep CNN method in [17] is used for face detection and recognition. The recognition system includes FaceNet CNN, Face landmarks and face embedding and SVM classifier with accuracy 95.02%.
Similarly in [18], proposes a video based FR system with 99% accuracy on 82 students, each having 32 face images.
Deep learning based on improved AlexNet CNN model [19] proposed a recognition rate of around more than 80%. It provides better feature extraction. WebFace data Set is used to improve training of network and test.
The paper proposes face hallucination (FH) and recognition system [20] using generative adversarial network in which face hallucination reconstruct a high resolution image from a low resolution image and recognition can be accurately done by the restricted image. The research proposes two models, first FH-GAN network which has improved face hallucination and recognition together. Second is a new hallucination network, dense sparse network which was implemented jointly with DSNet for better result. The dataset used in this proposed work are LFW for testing the performance of face hallucination and also for checking the accuracy of face recognition.

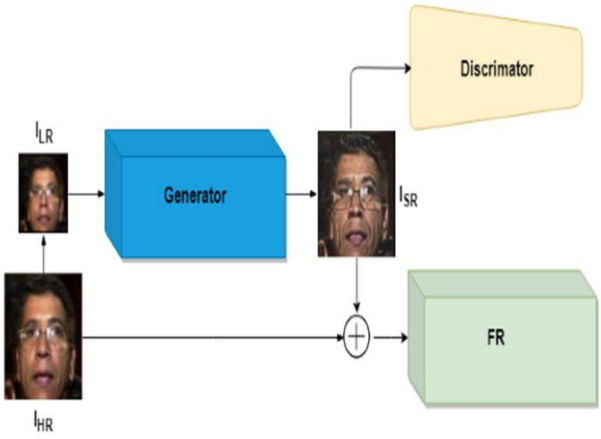


Fig4: FH-GAN architecture [20]

An algorithm for face authentication with unconstrained condition (UC), based on deep convolutional features is proposed and evaluation done in IARPA Janus Benchmark A (IJB-A) dataset and Wild (LFW) dataset in [21]. The IJB-A dataset includes 500 subjects with pose and illumination variations. The CASIA- WebFace dataset is used to train deep convolutional neural network (DCNN). Landmark detection is done before preprocessing, as it gives better performance on unconstrained faces which was followed by face alignment. Bayesian Metric learning is proposed for using positive and negative image in the training set. This has enhanced the performance of face verification (FV). For the evaluation of algorithm, cumulative match characteristics scores and receiver operating characteristic curve has been used.

E. Standard Dataset used in the reviewed paper with Key Features

Table 1: Standard dataset used [13]-[15]-[20]

Dataset	Image	Subjects (S)	Image(I) used	Key Features
[15] VGGface2	3.3 M	9000	1.2M	8631 S Pose, age, ethnicity for FR.
[20] CASIA-webFace	4,94414	10,575	Less than 500K	Celebrity
CFP[20]	7000	500	-	Used for FV
LFW[20]	13233	5749	-	Testing FV &FH
JANUS CS2 (IJB-A)[20]	5397 image,2042 videos(V)	500	5397 image,2042 videos(V)	V split into167 (gallery set) x 1806. (probe set)

MSceleb-1 M [15]	10M	100	-	Celebrity
[13] Yale database	165	15	10 image of each subject	11 image per subject for expression or lighting.
[13] ORL face database	10 for each subject	40	9	open& closed eyes, smiling and non-smiling face
FERET face database [13]	6/subject	70	5	Upright and front view faces

III. PROPOSED WORK

A.Face Recognition using Convolution Neural Network

The face recognition involves detection of the face from the image, computation of face embedding, training and recognizing the face. In the detection process the presence and location of face is obtained. Face landmark is done to locate the face and to detect the facial structure of the face. This is followed by embedding [16]- [17] done to extract the 128-d feature vectors from the face. Each face for training has three images. Training is done to optimize the face embedding process done in feature extraction.

The steps involved are

- Image Capture
- Face Detection
- Face landmark detection
- Face embedding
- Training
- Recognizer

Image generation for the dataset is implemented through webcam of laptop. 50 images are stored in a particular folder with name of the person. 5 persons dataset has been created. Face detection is performed with CNN cascade. Torch based deep learning method is used for detection. Face embedding is performed with FaceNet model. Face similarity is based on Euclidean space. The L2 distance between the faces determines the similarity. This is obtained by

$$d_{L2}(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

There are various methods to train the model. One of the methods is triplet loss function which minimizes the distance between positive and anchor or maximizes the distance between anchor and negative. The anchor which is the current face, the positive image is similar to given image and third is negative image, not similar to the current image. The neural network computes the embedding of each face and pulls the weights via triplet loss function so that positive image lie close to anchor and negative image are pushed away by maximizing the distance between them. The triplet loss function equation used for computation has function for encoding of anchor, positive and negative images. The triplet loss equation is used for calculating the loss, which is depicted below

$$= \sum_{i=1}^{N}\left[\left\| f_i^a - f_i^p \right\|_2^2 - \left\| f_i^a - \left\| f_i^n \right\|_2^2 + \alpha\right]$$

Where the function f(a) stands for output encoding of the anchor, f(p) for output encoding of positive and f(n) for output encoding of negative. α is a constant, so that the difference between anchor and positive and difference between anchor and negative are not zero [16].

The various parameters which are taken into account during landmark detection for better accuracy are pose, sharpness, resolution and brightness etc. Pose estimation of students can be calculated by using landmark face detection. The pose of the face may be different direction when the image is captured i.e. in different angles. The pose can be measured in three angles, roll, yaw and pitch. The equation for yaw angle for two points (x_1, x_2) and (y_1, y_2), is given as

$$\text{Yaw} = abs(\arctan 2(y_2 - y_1, x_2 - x_1))$$

When the image is captured from the video, then it is possible that some of the images may be blurred. Variance of image Laplacian is used to calculate the sharpness of image.

$$\text{Sharpness} = \sum_{(i,j)\varepsilon \upsilon(x,y)} \Delta I(i,j) - \overline{\Delta I})^2$$

Where $\overline{\Delta I}$ is denoting the mean value of image laplacian in $\upsilon(x, y)$. The position of eye corners are measured in face landmark and distance between the right eye corner and left eye corner is used for calculating the resolution.

$$\text{Resolution} = \sqrt{(x_L - x_R)^2 + (y_L - y_R)^2}$$

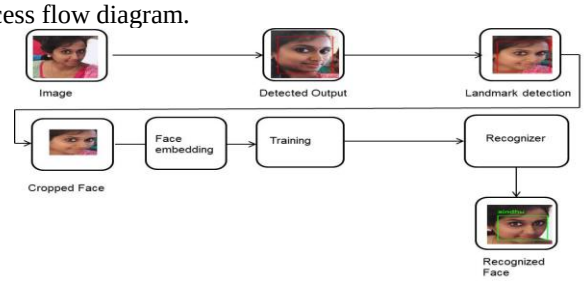
Where (x_L, y_L) and (x_R, y_R) are the left and right eye corners respectively. The brightness of the image is calculated by the mean of the red, green and blue channels intensities [22].

$$\text{Brightness} = (R + G + B)/3$$

Linear support vector machine (SVM) classification is used, which will transform the data for finding out the optimal boundary between the possible outputs. In the training part, SVM, machine learning model is trained to recognize the face. The details obtained in the extraction embedding, training is used in the recognizer to recognize the faces.

B. Model for Face Recognition

The steps involved in face Recognition are face detection, landmark detection, face embedding, training and SVM classifier. The illustration of result obtained is presented by process flow diagram.



C.Result

The following are result face detection & face recognition. The detection rate obtained is 100% for images tried.

Face recognition implemented correctly for all the four faces.

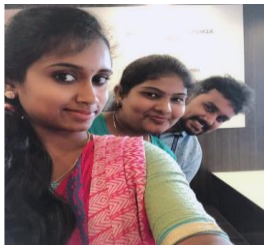



Fig 6: test images1 &2 selected for the detection

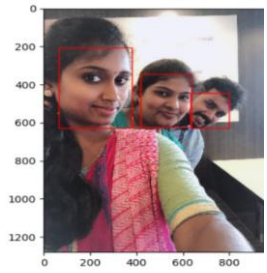


Fig 7: Detected faces in the Image

The presence and location of three faces has been obtained using CNN cascade.

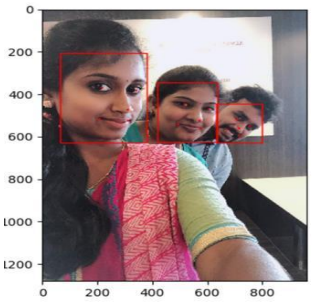


Fig 8: face landmark detection

Fig 9: Cropped faces from the image

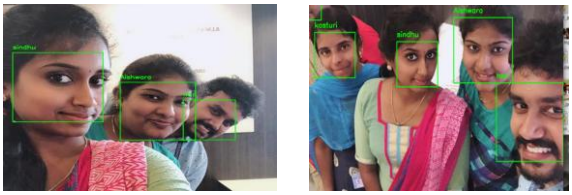


Fig 10: Face recognition of faces in the test images 1 &2.

Result of recognition using two test images & webcam. Dataset has been prepared manually, 10 images per person and then image generator using the python script.

Table2: test result from two images.

Mode of preparation of dataset	Image Dataset	Test Image No	No of face detected	No of face recognized
Manually	40 (10/person)	1	3/3	3/3
		2	4/4	4/4
Image Generator with webcam	200(50/ person)	2	4/4	4/4
Image Generator with webcam	200(50/ person)	-	4	4

IV. CONCLUSIONS

Different types of face recognition system have been reviewed in holistic, feature, hybrid and deep learning approach. Face recognition system using convolution neural network is implemented with face detection, landmark detection, face embedding, training in deep learning network. 100% detection rate obtained for face detection with CNN cascade. Face embedding using FaceNet CNN is done successfully to quantify the face. Face recognition has been implemented successfully and accurately for smaller dataset. As a future work, face recognition will be carried out with a larger dataset.

REFERENCES

1. Divyarajsinh, N.Parmar, Brijesh B. Mehta, " Face recognition methods and application" , International Journal of Computer and Technology and Application, vol 4(1), 84-86, ISSN:2229-6093,Jan-Feb 2013.
2. Refik Samet, Muhammed Tanrivendi, " Face recognition based mobile Automatic Attendance Management System", 2017 International Conference on Cyber worlds, DOI: 10.1109/CW.2017.34, Chester UK.
3. Chengjun Liu and Harry Wechsler, "Comparative Assessment of Independent Component Analysis (ICA) for face recognition", 2nd International conference on Audio and video Biometric person authentication, AVBPA'99, Washington D.C U.S.A, March 22-24, 1999.
4. Mathew Tunk and Alex Pentland, " Eigenfaces for recognition", Journal of Cognitive Neuroscience, volume 3, Issue 1,1991, p.71-86.
5. Mustamin Anggo, La Arapu, "Face Recognition using Fisher face method" Journal in Physics conference series, 1028(1): 012119, June 2018, DOI: 10.1088/1742-6596/1028/1/012119.
6. Omar Abdul Rhman Salim, Rashidah Funke Olanrewaju, Wasiu Adebayo Balogum"Class Attendance Management System using Face Recognition", 7th international Conference on Computer & communication Engineering, September 2018, DOI: 10.1109/ICCE.2018.8539274.
7. Priyanka Wagh, Roshani Thakare, Jagruti Chaudhari, Shweta patil, " Attendance System based on Face Recognition using Eigen face and PCA algorithm". 2015, International conference on Green Computing and Internet of Things, 8-10 October 2015, DOI: 10.1109/ICGCIoT.2015.7380478.
8. Takeo Kanade, Dept. of Information Science, Kyoto University, 1973, PhD Thesis report "Picture processing system by computer complex and recognition of human face".
9. Laurenz Wiskott, Jean Marc Fellous, Norbert Kruger and Christoph Von der Malsberg, "Face recognition by Elastic bunch graph matching", in Intelligent Biometric Techniques in Finger Print and Face Recognition, eds.L.C Jain et at.,pub1.CRC Press, ISBN 0-8493-2055-0, Chapter 11, pp.. 355-396,1999.

10. Trupti M. Kodinariya, Dr. Prasant R. "Hybrid Approach to face recognition system using principal component analysis and Independent Component Analysis with score based performance", CoRR abs/1401.0395(2014).
11. Ushamary Sharma, " Hybrid feature based face verification & recognition system using PCA and Artificial Neural Network", Indian Journal of Science and Technology, vol:8(S1), 115-120, February 2015, DOI:10.17485/ijst/2015/v8iS1/57754.
12. Aman Dhiman, Raman Kumar, " A Hybrid Approach for the detection and recognition of faces with PCA and enhanced K- mean clustering in facial image database", International Journal of Engineering Research and General Science, vol 5, issue2, March-April 2017, ISSN 20191-2730.
13. Rui Huang, Vladimir Pavlovic, Dimitris N. Metaxas, " A Hybrid Face Recognition Method using Markov Random Fields", 17th International Conference on pattern recognition(ICPR), Cambridge, UK, 1051-4651/04 IEEE, ISBN 0-7695-2128-2, DOI: 10.1109/ICPR.2004.1334492.
14. Patrik Kamencay, Miroslav Benco, Jomas Mizdos, Roman Radil, " A new method for face recognition using Convolution Neural Network", Digital image processing & Computer graphics, volume 5, No. 19, 2017, Special issue, Advances in Electrical and Electronics Engineering, DOI: 10.15598/aeee.v15i4.2389.
15. Qiong Cao, Li Shen, Weidi Xie, Omkar M. Parkhi and Andrew Zisserman, " VGG face2: A dataset for recognizing faces across pose and age", accepted in 2018 13th IEEE conference on Automatic Face and Gesture Recognition(F& G), May 15-18 2018,.
16. Florian Schroff, Dmitry Kalenichenko, James Philbin, "FaceNet: A unified embedding for face recognition", arXv: 1503.03832v3[cs.CV], 17th June 2015.
17. Marko Arsenovic, Srdjan Sladojevic, Andras Anderla, Darko Stefanovic, "Face Time-Deep Learning Based Recognition Attendance System", 15th international symposium on Intelligent System and Informatics, Sept. 14-16, 2017, Subotica, Serbia, DOI: 10.1109/SISY.2017.8080587.
18. Lin Zhi-Heng, Li Yong-zhen, "Design and Implementation of Classroom Attendance System Based on Video Face Recognition", 2019 international Conference on Intelligent Transportation, Big Data & Smart City(ICITBS), Changsha,China,2019,pp.385-388. DOI: 10.1109/ICITBS.2019.00101.
19. Wenxian .Zeng, Qinglin Meng, Ran Li, "Design of Intelligent Classroom Attendance System Based on Face Recognition", 2019 IEEE 3rd International Technology, Networking Electronics and Automation Control Conference(ITNEC), Chengdu,China, 2019, pp.611-615, doi: 10.1109/ITNEC.2019.8729496.
20. Bayram Bayramli, Usman Ali, Te Qi, Hongtao Lu, "FH GAN: Face Hallucination and Recognition using Generative adversarial network", arXiv: 1905.06537v1[cs.CV], 16th May 2019.
21. Jun-Cheng Chen, Vishal M. Patel, and Rama Chellappa, " Unconstrained Face Verification using Deep CNN features", 2016 winter IEEE conference on application of Computer Vision, 7-10 March, Lak Placid, NY, USA, DOI: 10.1109/WACV.2016.7477557.
22. Shubhobrata Bhattacharya, Gowtham Sandeep Nainala, Prosenjit Das, Aurobinda Routray, "Smart Attendance Monitoring System (SAMS): A Face Recognition based Attendance System for classroom environment", 2018 IEEE 18th International conference on Advanced Learning Technologies,9-13 July 2018, Mumbai, India, DOI: 10.1109/ICALT.2018.00090..

AUTHORS PROFILE

Dr.R.S.Sabeenian is currently working as a Professor and Head in ECE Department in Sona College of Technology, Salem, Tamil Nadu, India. He has more than 18 years of teaching experience. He has published more than 100 research papers in various International, National Journals and Conferences. He has received 10 National level awards for his outstanding research contribution and academic excellence from IETE, ISTE, IEEE, Springer, Nehru Groups, INAE, ISTE TP, TNSCST and CICR. He has also received Project Grants of more than 3.5 Crore from AICTE, DST and reputed industry. He also received International funded project worth more than 15 lakhs from Nanolytix, Canada, Consultancy project from Titan worth Rs.10 lakhs. He has produced 10 doctorates (Ph.D Candidates).He has filed 3 patents and one of his patent granted in the year 2019. He is the Editor of 6 International Research Journals. He is also associated with the Image Processing Payload of the PESIT Pico Satellite Project. He is the Expert Member for Board of Studies of more than five autonomous institutions. He delivered more than 70 guest lectures and chaired more than 25 national and international conferences. He has also Received "AICTE Career Award for Young Teacher (CAYT)" from AICTE for the period of 2015-18.

J.Harirajkumar, Associate Professor, Department of ECE, Sona College of Technology, Salem has an academic experience of 12 years, he has specialized in VLSI Design and Nano electronics. He is pursuing his Ph.D in the area of VLSI Design under Anna University, Chennai. He is a member of IEEE and ISTE. He has published 5 National and International journals.

Lizzie D' Cruz is working as Lecturer (SG) in Dr. B. R. Ambedkar Institute of technology, Port Blair, Andaman & Nicobar Islands, has an academic experience of 28 years. She is pursuing her M.E in VLSI Design from Sona College of Technology, Salem, Tamil Nadu, India. She has published four papers in international journals. She has been awarded Best Engineering Practice adopted in the field of Engineering Education in the year 2008.