Virtual ATM through Fingerprint and Face recognition using Deep Learning

Vaishnavi M. Shinde[1], Dhanashri B. Shinde[2], Pravin B. Shinde[3], Aditya V. Wandhekar[4], Pallavi S. Kohakade[5]
[1,2,3,4,5]*Computer Dept., SCSMCOE, Ahmednagar*

Abstract **Fingerprints and facial features of the individual are being used in biometric authentication techniques, which are increasingly extensively used across significant implementations. Despite the fact that there multiple facial recognition systems accessible. A greater number of research should unearth factors that improve efficiency and accuracy. Facial as well as fingerprint identification play an important part in the identifying process since they do not need human assistance, unlike some other biometrics methods. This not only proves the huge potential to create far greater protection for such Virtual ATM transactions, but also explains the reasoning why biometric identification systems have been attracting so much attention. Therefore, for this purpose an effective framework for biometric authentication on Virtual ATMs through the use of biometric features, such as Facial and Fingerprint have been proposed. The presented framework utilizes Live Streaming and Region of Interest along with Channel boosted Convolutional Neural Networks and OTP authentication has been implemented. The framework has been measured using lengthy experimentations to achieve quite reassuring outcomes.**

Keywords: Virtual ATM, Biometric recognition. Face recognition, Fingerprint Recognition, Channel Boosted Convolutional Neural Networks.

I INTRODUCTION

Internet has arrived since both the Internet and smartphones devices have become more commonplace. Here has been a meteoric rise in the usage of smartphones over the last several decades, and with that rise has come an influx of new smartphone applications and software programmers. Here, a biometrics ATM technology using machine vision and technology that recognizes faces is a game-changer for productivity. The person's face is a good base for personal identification and verification due to its substantial individual traits and self-stability as a characteristic physiological component intrinsic to the human body. When contrasted with other approaches, facial recognition is universally welcomed, pleasant, trustworthy, and risk-free.

Numerous different types of equipment have been introduced in India as a result of technical advancement, each with the intention of raising customer satisfaction. The automatic teller machine (ATM) was one item that streamlined banking for financial institutions. The advent of ATMs enabled customers to conduct financial transactions independently. Originally, only customers of a certain bank could utilize an ATM to move money; later, however, all ATMs became connected to a single network, enabling customers of just about any bank to utilize any ATM of their choosing. Due to this, customers of other banks remained able to use the ATMs of other banks to make deposits, withdrawals, and wire transfers.

Facial detection is so natural to human people that even newborns can tell family and friends apart. However, computers have a hard time with face detection. In the inaugural automated face identification system, an extracted features is built by labelling the locations of facial landmarks like the eyeballs, eyebrows, chin, etc., and faces are identified by measuring the Euclidean distance between extracted features from different photos. In order to characterize the architectural aspects of facial photographs in a big database, most methods employ feature maps of varying dimensions. While others characteristic face recognition methods simplify the categorization work by treating the face region as a juncture and representing it in a lower dimensions environment that is generated from the multidimensional space input images.

Facial identification is considered one of the most reliable methods of establishing a person's identification, which has been the topic of a great deal of research over the last numerous years. Face recognition from photographs is a popular area of study in biometrics. Among the most useful applications of facial recognition and identification technologies is evaluating images for interpretation. Mental health professionals, neuroscientists, and machine learning experts have all taken an interest in facial recognition software since advances in the field may provide light regarding how the natural brain processes. Despite the prevalence of biometric identification methods like fingerprint as well as retinal scanners, they still rely on human verification. Nevertheless, it is not necessary while utilizing face pictures for person authentication. Facial recognition technology plays a crucial part in establishing a person's identity since it does not truly entail human cooperation that is a considerable benefit above all other biometrics approaches.

According to research by Chaoyou Fu [1], a novel Double Variational Regeneration Face architecture have been designed to enhance the efficacy of Heterogeneous Face Recognition by generating large numbers of paired diverse images from playback. The very first stage is painstakingly building a double variational encoder that really can train with both associated huge datasets and imbalanced visible information. The addition of this feature greatly broadens the scope of possible identifications from the resulting images. A bilateral identification conservation decline is then applied to the generated images to guarantee that they retain their original consistency. This new set of unidentified images may be utilized to train up the Heterogeneous Face Recognition models using descriptive learning, capitalizing on both the identification persistence and identification dispersion features.

Using face recognition technology, Kanjana Eiamsaard [2] developed the Smart Warehouse Accessibility Management System. Smart Warehouse Accessibility Monitoring Program was validated using a confusion matrix experimental process. The system as a whole performed as expected in terms of accessible surveillance. Crime investigations may be sped up and made more efficient with the help of the Smart Warehouse Accessibility Surveillance System. Upgrading the Smart Warehouse Accessibility Management System to include item identification might allow for constant tracking of stolen goods and immediate notification in the event of a breach.

[3] Feng Liu Explain a novel regression-based approach to face identification and face recognition reconstitution from a single 2D image, allowing for any expression or pose. It takes a 2D image of a face and uses those cues as clues to rebuild the 3D face, subsequently utilizing those revised 3D faces to further enhance the 2D characteristics. In order to do both tasks simultaneously in real time, the proposed method alternates between employing feedforward landmark regression model and 3D shape coefficient of determination. The proposed approach outperforms state-of-the-art 3D face remediation tools by automatically recreating both pose- and affirmation and expressive 3D shapes from a single face photo of arbitrary postures and expressions.

Section 2 of this research article presents an analysis of the relevant literature; Section 3 explains the research approach; Section 4 discusses the experimental assessments; and Section 5 closes with suggestions for further study in the future.

II RELATED WORKS

Since there is a limit to the number of Eigen faces that can be used in Principal Component Analysis transformation, Gurlove Singh [4] reports that the method was not more successful than other techniques, which includes traditional and digitalized face recognition. More work is needed to perfect the fully automated front-view facial recognition technology that shows off pinpoint accuracy in demos. There is going to be substantial improvement in the system's precision when used in practice. To achieve a high degree of precision, the mechanism was badly conceived and built. One contributing factor would be that the face detection and recognition platform's component is not always sensitive enough to subtle changes in consistency with respect to dimensionality or orientation.

An end-to-end learning strategy for pose-invariant facial expression recognition and face picture reconstruction using geometric information is described by Feifei Zhang [5]. In order to help in the training of a deep neural binary classifier, it may generate face images with fabricated expressions and postures. Extensive experiments on three most commonly used datasets show the effectiveness of this method. Facial features may also be transmitted using this method.

According to Zhang Jianxin [6], detecting face traits for identification has been among the most difficult tasks. This article's authors offer a novel technique for recognizing people's faces by using the Two-Dimensional Adaptive Directional Wavelet Transform and indeed the Latent Semantic Feature space Methodology. Therefore, for the very first instance, researchers apply the adaptable directional wavelet transform to the face identification issue in two dimensions. Two-dimensional adaptable directed wavelet decomposition is used to make an iterative forecasting and refresh step that takes into account the local features connecting individual pixels. Both of the forecasting and refresh controls in this study use a set of nine orientations with 8 fading occurrences. The ideal orientation for performing the hoisting wavelet decomposition is determined by applying a thread segmentation technique to the data.

Using Eigen Face, fisher face, and Local Binary pattern approaches from the Open source domain, Limei Fu [7] offers a system for recognizing and identifying people's faces on Linux. The minimization strategies are therefore put out via a validation process, and investigation is utilized to analyses their pros and cons as well as the conditions in where they could be applied. The development of reliable software that recognizes faces is a complex issue. Without using many methods, achieving a respectable identification outcome is difficult. By merging local and worldwide data, facial features may be accurately characterized. Increasing identification effectiveness in tandem with the use of numerous features and classifiers may be achieved via the use of the strategy of combining both.

According to Sergei Shavetov's [8] study, the suggested method is low-cost, user-friendly, and uncomplicated to implement while providing just the basic minimum of protection versus malicious actors. While it may function well in fully automated operation, human supervision throughout the verification process is recommended to prevent any unplanned bypasses of the safeguards. It might be used as a certificate authority, in conjunction with the likes of radio-frequency identification, near-field interaction, and other methods now in use. If the genuine identifier is misplaced or misappropriated, the program will be absolutely certain that the wrong person is prevented from entering, independent of how well the alternative type of validation works.

In [9], Di Wang explains how convolutional neural networks are an important component of reinforcement learning. Particular advantages in computer vision may be derived by using convolutional neural networks, which are based on local circuit associated parameters and other properties. Development at Convolutional Neural Networks is vital since it affects both the efficiency in which the system is trained and the final efficiency of recognition. Drop - out rates value, activation function, and number of modeling layers are investigated as design approaches. Improved identification accuracy is the outcome of utilizing the original data set to refine the strategy. Given how difficult it is to collect a significant amount of information in practice, the data set utilized in this study has constraints, but it could nevertheless be utilized to help prove the efficacy of the novel process.

According to Bharath Tej Chinimilli [10], Local Binary Pattern Histogram is one of the most well-known methods of face identification. This method is useful for picking out pupils who have accidentally changed their appearance (for example, by wearing glasses or growing facial hair). The small size of the dataset raises certain concerns. It is possible that a new, improved data set will be compiled in an effort to get a more reliable result. The authors potentially improve the Haar cascade algorithms' capacity to recognize new faces if they generate a new training set. The program may sound an audible or visual warning if a cheater is detected in the classroom.

A novel graph-based, multi-Face Enhancement Generative Adversarial Network has been released by Mandi Luo [11], expanding existing datasets for deformation-invariant facial recognition software. Instead of only isolating the personality prognostications, it also utilizes the dissociated representations to alter face attributes, which greatly improves the reliability of biometric identification programs. The authors furthermore provide Graph Convolutional Systems for investigating global connections amongst different face regions, which helps to better preserve the spatial information. Extensive experiments on face recognition and picture reconstruction tasks show that our proposed method is successful in learning a superior identity-preserving capacity from restricted datasets.

To aid with this endeavor, Gou Wei [12] proposes a multi-feature structure for the sentiment classification issue. There are two parts to the data on emotions: video and audio. Traditional technologies is used to extract meaningful from auditory, and 4 main channels are used to retrieve sentiment traits from video sequences, all with the help of the authors. A weighted sum is then used to incorporate the test score, with each network's weight based on how dedicated it is to the best possible result. Researchers found that the conceptual approach performed better than the competition in experimental evaluations of perceptual processing in the environment.

M. Geetha [13] recommends using a machine learning-based facial identification approach with a support vector machines framework to keep an eye on students' online test activities. The proposed method aids in expedited face identification by synthesizing feature maps from image features. A variety of other techniques, such as the Speeded up Robust Functionalities, the Scale Invariant Feature Transform, the Fisher faces, and the binary pattern local histograms, may indeed be utilized to develop more resilient object recognition which can locate faces amid varying lighting circumstances. Implementing other approaches might lead to superior optimum values as well. To improve accuracy, convolutional neural networks may be used.

III PROPOSED METHODOLGY

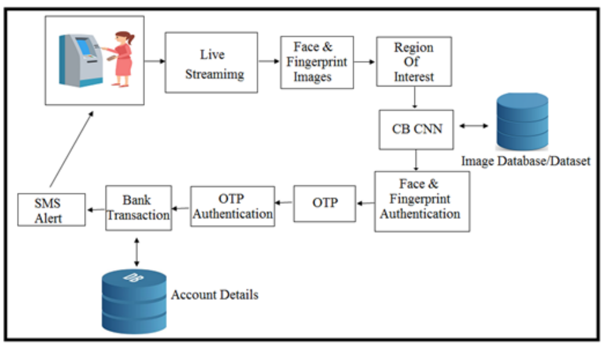

Figure 1: Proposed Methodology

The proposed approach for achieving Card less Virtual ATM system that utilizes biometric authentication in the form of face and fingerprint authentication has been depicted in the figure 1 above and the steps taken to achieve this system are elaborated below.

Step 1: GUI Building – For the purpose of demonstration of our methodology, an interactive user interface has been realized using the swings framework on the java programming language. The application for performing bank transactions by the customer has been achieved as an interface that is easy to understand and navigate. The developed and designed interface facilitates the enrollment of the users into the platform by collecting their various attributes and information such as name, date of birth, mobile number etc. along with the user authentication components such as face and fingerprints. The user characteristics such as facial characteristics captured through the facial image of the user taken at the time of registration are stored along with the fingerprints of the user. These characteristics are useful for the realization of the virtual ATM approach that will verify the users based on the facial features and fingerprints. The collected characteristics of the specific users are segregated and stored in an image database that will be utilized in the further steps for the purpose of training our deep learning model.

Step 2: Training the CB-CNN model – The collected images from the user having the users face and fingerprints respectively have been crucial in the realization of the training of the deep learning model selected for this implementation. The model being utilized for the purpose of achieving the training and recognition of the user through the face and fingerprints is the CB –CNN model.

The CB-CNN stands for Channel Boosted Convolutional Neural Networks which are an improvement over the traditional Convolutional Neural Networks. The CB-CNN model utilized boosted color channels for the implementation over the conventional CNN that does not incorporate any boosted channels. The CB-CNN approach can considerably improve the recognition accuracy in comparison to the traditional CNN due to the implementation of the boosted channels.

The Sequential class incorporated into the TensorFlow library allows for the development of sequential neural network architectures. Next, as the first stage of the CB-CNN Design, we add a convolution layer with 32 3x3 kernels and the ReLU activation function, reserving this layer for images of the correct size. It is the only purpose of this layer to ensure that the images are the same size. Afterwards, a Convolution layer with 64 3 x 3 kernels and ReLU activation is introduced. A maxpooling layer with a dropout regularity of 25% and a size of 2 by 2 units has been planned.

Each of the 128 kernels has a size of 3 by 3, owing to the extra fully connected layers. To do this, we have used a special activation function called the ReLU activation function. The maximum pooling layer now has the prescribed size of 2x2. After the third layer is complete, the last layer is implemented with 128 3x3 kernels and the ReLU activation function. We add a second Max pooling layer, adjusting the dropout to 25% and keeping the dimensions at 2 x 2.

After the neural network training is complete, it is flattened using the flatten method, a dense layer of size 1024, and the ReLU activation function. A dropout ratio of 50 is needed at the conclusion of the convolution neural network. Following which a dense layer has been implemented with 7 classes, each of which corresponds with the 7 users that are being utilized for the demonstration of this methodology.

It is common procedure to utilize the Adam optimizer to improve the outcome with 500 epochs for all of the user attributes, including their face and fingerprints, while the player remains within the learning phase. When the training stage is finished, the model imports the learned data from an H5 file and uses it during the testing phase. The structure of the Channel Boosted – Convolutional Neural Network is shown in Figure 2.

Layer	Activation
CONV 2D 32 X 3 X 3	Relu
CONV 2D 64 X 3 X 3	Relu
MaxPooling2D 2 X 2	
Dropout 0.25	
CONV 2D 128 X 3 X 3	Relu
MaxPooling2D 2 X 2	
CONV 2D 128 X 3 X 3	Relu
MaxPooling2D 2 X 2	
Dropout 0.25	
Flatten	
Dense 1024	Relu
Dropout 0.25	
Dense 7	Softmax
Adam Optimizer	

Figure 2: Convolution Neural Network Architecture

Step 3: User Authentication and Transaction Completion – The trained model achieved in the previous step is being utilized for the purpose of achieving the authentication of the user for completing the transaction. The registered user interacts with the virtual ATM to provide with the respective details such as face and fingerprint for the purpose of verification. The system captures the face and fingerprint images and provides to the subsequent modules for the purpose of authentication. The images are utilized by system by first subjecting the images to the process of region of interest evaluation. The region of interest isolates the region of the face as well as the fingerprints for the evaluation by the trained model. The images are then subjected to the trained model that verifies the user based on the captured facial and fingerprint images.

Once the user has been verified a One Time Password or an OTP is generated and sent to the user via email. This OTP needs to be authenticated before performing the transaction. As the OTP is authenticated the bank transaction initiates and the system navigates the user to perform the bank transaction efficiently.

IV RESULTS AND DISCUSSIONS

The proposed approach for the virtual ATM through the use of facial as well as fingerprint features of the user. Both the Python and Java programming languages, using the Spyder and NetBeans integrated development environments, respectively, have already adopted the concept. The code was written in both of these contexts. The plan requires using the OpenCV, TensorFlow, and Keras frameworks to do the necessary tasks of deep learning. It was determined that a laptop with an Intel Core i5 CPU, 8 GB of Memory, and 1 TB of storage space worked best for testing the forthcoming implementation.

Accuracy in face and fingerprint detection must be measured with the efficiency of the proposed technique. Error may be used to gauge how trustworthy a technique is; typically, the lower the error, the more trustworthy it will be. A reliable error analysis may be carried out with the use of the root mean square error measure.

Performance Evaluation based on RMSE

The Root Mean Square Error is one of the most useful performance measures for assessing the degree of deviation between a set of similar features. The suggested method will be evaluated based on two metrics: the percentage of successfully identified fingerprints and the percentage of incorrectly identified fingerprints. Here, Equation 1 is given for your convenience in calculating the RMSE.

$$\mathbf{RMSE}_{fo} = [\sum_{i=1}^{N} (z_{f_i} - z_{o_i})^2 / N]^{1/2}$$

Where,

$\sum$ - Summation

$(Z_{fi} - Z_{oi})^2$ - Differences Squared for the fingerprints identified correctly and fingerprints identified incorrectly

N - Number of conducted Experiments.

For various iterations of fingerprint recognition using the proposed technique, we compute the relative mean square error or RMSE values. Fingerprint recognition is tested ten times to ensure accuracy. The proposed method's identification result is saved after each session it is acquired. Then, the information is used to carry out an RMSE evaluation. Careful calculation of the outcomes described in table 1 that would be presented below yields these RMSE readings.

S no.	Number of Iterations	Correctly identified Fingerprints	Incorrectly identified Fingerprints	MSE
1	10	9	1	1
2	10	8	2	4
3	10	9	1	1
4	10	7	3	9
5	10	10	0	0

Table 1: RMSE measurements tabulated.

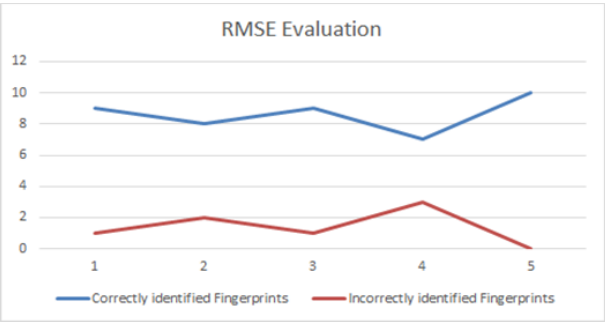

Figure 3: Line Graph for RMSE outcomes

For the purpose of making the diagram in the figure 3 supplied above, we are using the results acquired for the identification effectiveness and the RMSE measurements in the table 1 previously provided. The following table and graph illustrate the method that is meant for the purpose of fingerprint recognition in order to obtain a very small error rate of 1.732.

Performance Evaluation based on Precision and Recall

Measures of precision and recall are very useful for assessing the thoroughness with which a certain module of the paradigm has been executed. These two measurements are discussed in the larger framework of our approach. Relative correctness is defined by the module's precision, which includes its reliability across a wide range.

The accuracy of this approach was measured by comparing the number of correct identifications with the total number of tests. Nonetheless, the recall requirements are a useful adjunct to the accuracy assessment in determining the CB-CNN component's overall dependability. This is due to the fact that precision monitoring is not adequate on its own.

The recall is determined in this method by contrasting the proportion of correct to incorrect identifications. In order to quantify this argument, the following equations are provided.

Precision and Recall can be depicted as below:

- ✓ A = The number of accurate face identifications
- ✓ B = The number of inaccurate face identifications
- ✓ C = The number of accurate face identifications not done

So, precision can be defined as

Precision = (A / (A+ B)) *100

Recall = (B / (B+ C)) *100

Using the abovementioned formula, the experimental findings are shown in Table 2 below. With these statistical parameters, we can generate the representation illustrated in figure 4.

Table 2: Precision and Recall Measurement Table

No. of Iterations	Accurate Face identifications (X)	Inaccurate Face identifications (Y)	Accurate Face Identifications not done (Z)	Precision	Recall
10	10	0	0	100	100
20	17	2	2	89.47368	89.47368
30	27	2	2	93.10345	93.10345
40	35	3	3	92.10526	92.10526
50	44	5	2	89.79592	95.65217

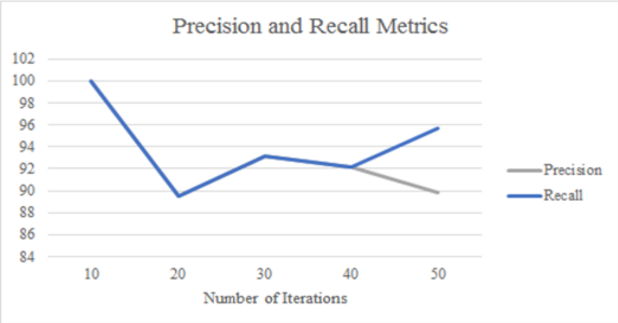

Figure 4: Comparison of Precision and Recall

In this graph, we see the CB-CNN in action, demonstrating its capacity to perform over a broad variety of trial counts and provide accurate face detection in accordance with the input data. The very high accuracy and recall rates of 92.89 and 94.06 percent, respectively, demonstrate the method's reliability. These numbers are quite sizable for the first implementation of such a technique, and the success that has resulted is commendable.

V CONCLUSION AND FUTURE SCOPE

The proposed approach for achieving Card less Virtual ATM system that utilizes biometric authentication in the form of face and fingerprint authentication has been elongated in this research article. The presented approach utilizes the ATM machine along with the camera for capturing the live feed of the user. The live feed from the ATM camera is being streamed to the system live which is then effectively utilized for the purpose of achieving face and fingerprints of the user. These are converted into images and provided to the next step for the region of interest estimation. The region of interest isolates the facial features as well as the region of the image containing the face. These are provided to the Channel Boosted Convolutional Neural networks that is connected to the database consisting of user facial and fingerprint images. The CB-CNN approach then effectively authenticates the user based on the images in the database and a One Time Password is sent to the registered user mobile number. This OTP is then entered by the user which is authenticated and the user is allowed to perform the bank transaction on the ATM and a suitable alert is sent to the user via an SMS. The approach has been rigorously evaluated for its performance using RMSE and Precision and Recall which has resulted in highly reassuring measurements.

In the future this can be implemented using the Reinforcement learning model to recognize the aging face accurately.

REFERENCE

[1] C. Fu, X. Wu, Y. Hu, H. Huang and R. He, "DVG-Face: Dual Variational Generation for Heterogeneous Face Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 6, pp. 2938-2952, 1 June 2022, doi: 10.1109/TPAMI.2021.3052549.

[2] K. Eiamsaard, P. Bamrungthai and S. Jitpakdeebodin, "Smart Inventory Access Monitoring System (SIAMS) using Embedded System with Face Recognition," 2021 18th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2021, pp. 1-4, doi: 10.1109/JCSSE53117.2021.9493815.

[3] F. Liu, Q. Zhao, X. Liu and D. Zeng, "Joint Face Alignment and 3D Face Reconstruction with Application to Face Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 3, pp. 664-678, 1 March 2020, doi: 10.1109/TPAMI.2018.2885995.

[4] G. Singh and A. K. Goel, "Face Detection and Recognition System using Digital Image Processing," 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), 2020, pp. 348-352, doi:10.1109/ICIMIA48430.2020.9074838.

[5] F. Zhang, T. Zhang, Q. Mao and C. Xu, "Geometry Guided Pose-Invariant Facial Expression Recognition," in IEEE Transactions on Image Processing, vol. 29, pp. 4445-4460, 2020, doi: 10.1109/TIP.2020.2972114.

[6] Z. Jianxin and L. Haoran, "Local occluded face recognition based on 2D-DWT and sparse representation," 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Harbin, China, 2020, pp. 2110-2114, doi: 10.1109/ICMCCE51767.2020.00459.

[7] L. Fu and X. Shao, "Research and Implementation of Face Detection, Tracking and Recognition Based on Video," 2020 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), 2020, pp. 914-917, doi: 10.1109/ICITBS49701.2020.00202

[8] S. Shavetov and V. Sivtsov, "Access Control System Based on Face Recognition," 2020 7th International Conference on Control, Decision and Information Technologies (CoDIT), 2020, pp. 952-956, doi: 10.1109/CoDIT49905.2020.9263894.

[9] D. Wang, H. Yu, D. Wang and G. Li, "Face Recognition System Based on CNN," 2020 International Conference on Computer Information and Big Data Applications (CIBDA), 2020, pp. 470-473, doi: 10.1109/CIBDA50819.2020.00111.

[10] B. Tej Chinimilli, A. T., A. Kotturi, V. Reddy Kaipu and J. Varma Mandapati, "Face Recognition based Attendance System using Haar Cascade and Local Binary Pattern Histogram Algorithm," 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), 2020, pp. 701-704, doi: 10.1109/ICOEI48184.2020.9143046.

[11] M. Luo, J. Cao, X. Ma, X. Zhang and R. He, "FA-GAN: Face Augmentation GAN for Deformation-Invariant Face Recognition," in IEEE Transactions on Information Forensics and Security, vol. 16, pp. 2341-2355, 2021, doi: 10.1109/TIFS.2021.3053460.

[12] G. Wei, L. Jian and S. Mo, "Multimodal (Audio, Facial and Gesture) based Emotion Recognition challenge," 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), 2020, pp. 908-911, doi: 10.1109/FG47880.2020.00142.

[13] M. Geetha, R. S. Latha, S. K. Nivetha, S. Hariprasath, S. Gowtham and C. S. Deepak, "Design of face detection and recognition system to monitor students during online examinations using Machine Learning algorithms," 2021 International Conference on Computer Communication and Informatics (ICCCI), 2021, pp. 1-4, doi: 10.1109/ICCCI50826.2021.9402553.