Real-Time Face Recognition for Organisational Attendance Systems

Divyagna Bavikadi[1], Manjunatha A[1], Abhishek Pol[2], Akshat Kadam[3], Prajakta Kulkarni[3], Aparna Singh[3], P.M. Kamble[1], and Ravindra Hegadi[1]

[1] Central University of Karnataka, Kalaburgi, Karnataka, India - 585367
[2] Walchand College of Engineering, Sangli, Maharashtra, India - 416415
[3] Develearn Technologies Pvt Ltd, Mumbai, Maharashtra, India - 400019

Abstract. Today we are in an era of feasible biometric solutions to the age-old problem of verifying personal identity. As a form of identity that is verified through inseparable and unique characteristics of a person: their face - facial recognition is now one of the most popular methods in use. We propose an automated real-time facial attendance system where the users can verify their identity without physical contact with any surface. This is achieved through a conjunction of a mobile and platform-independent web application over a shared cloud database. The system is powered by a Face Recognition module to authenticate the users and demonstrated 99.7% of test accuracy, with an improved true positivity rate of 96.14% compared to some existing literature. Our system processes and aligns input face images before utilizing a deep convolutional neural network model to recognize the user's identity. To test the efficacy of our system, we have built a database of over 800 unique individuals of Indian descent. Our tests showed a superior true-positivity rate on our tweaked model demonstrating its efficacy in comparison with the reference literature. We also share some of the methods employed to raise system redundancy and minimize false negativity; a vital metric in any authentication application.

Keywords: deep convolutional neural network · face recognition · attendance system.

1 Introduction

Artificial Intelligence is rapidly updating the world and notably boosting the state-of-the-art in many applications including healthcare, security, academics, marketing, etc. In the sense of the attendance system, each organization has its unique process. At present, most of them do it manually by using the paper or file-based approach while few of them are marking automatic attendance using some biometric techniques.

Generally, all biometric systems begin with an enrolment process in which a person's unique characteristics are registered, followed by identification procedures. These two phases verify a specific user's biometric feature to a formerly stored

template taken upon enrolment. The templates for biometrics can be of many types like Fingerprints, Hand Geometry, Voice, Signature, Face, Gait, and Iris, however, the key authentications remain the same across all modalities. However, facial recognition has its own set of issues, finding it challenging to execute. Influences like variation in pose, facial hair, lighting based on the face location, background in the image, and facial or emotional expressions affect the image, and the result can vary accordingly. The face may not be detected in instances where the face is not visible or obscured from the camera. The majority of the existing systems have mobility, cost, security, accessibility, authenticity restrictions. Users must wait a long time in line while entering the organization via these techniques and involves the user giving actions that require physical contact with the external equipment.
An endeavor to conquer the inadequacies of the existing systems leads to the building of a Real-Time Face Recognition for Organisational Attendance System which will especially be useful in situations or future pandemic situations.

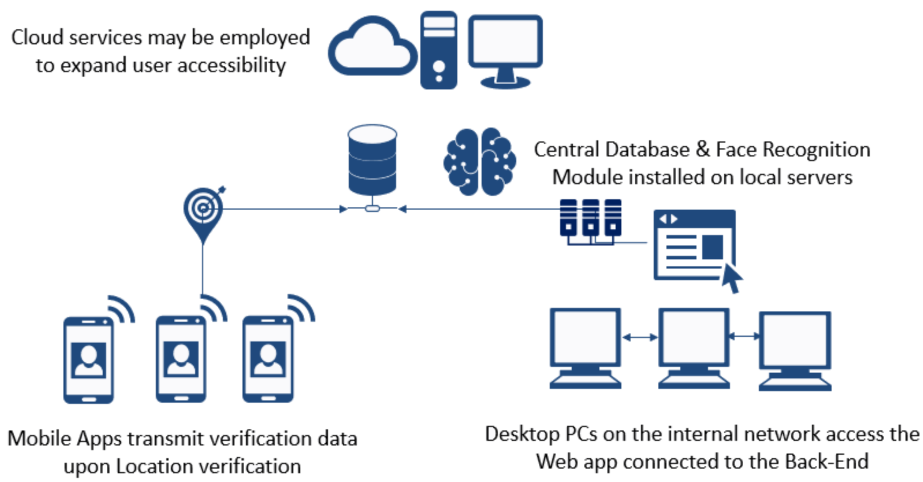

Fig. 1. Cross-device system design

Real-Time Face Recognition for Organisational Attendance Systems introduces a standalone application that includes a built-in responsive user interface with both web-based and mobile app-based versions, to make the attendance procedure seamless.
A note on the system design, it contains access levels that represent the end user's permissions and privileges across the system or within the corresponding class. The front-end access is given to all the users, though the dashboard and usable features have limited scope for basic/staff level users. The users have limited access to the database server depending on their respective access levels.

The Super Admin governs all the users including editing and deleting user details from the system. The Admin-level access has extended permissions and full authority over features and user members within the respective department unit. Admins will govern the staff-level users whose usage of the program is limited to Marking attendance, examining their details and attendance log. The access levels control tool lowers the danger of a privacy breach while also tackling a wide range of security concerns. The architecture mainly consists of three layers, an application layer, a server layer, and a backend database framework.
This work has the following contributions:
•Information Architecture of the system and how a web + mobile app works in tandem to fulfill base requirements.
•Unique checks and balances:
-Intranet-connection-based verification layer for Web app version.
-The local network (wi-fi) + geolocation-based verification layer for the mobile app version.
•Algorithm: How Facenet architecture was used to create a database-query protocol and how TP maximization/FP minimization was attempted by using multiple face vectors in a decision tree format rather than tweaking the architecture of the network. Fine tuning the similarity threshold with respect to the custom dataset created.
• There is no need for any additional hardware. User doesn't have to come in contact with any external equipment to mark attendance.

Paper organization: In section 2, literature is discussed. Section 3 provides the methodology of the end-to-end developed novel method and its architecture and working. Section 4 provides details about the testing phase and experiments done on the system, their results, discussion on experimental findings, the efficacy and robustness of the suggested strategy and architecture. Section 5 concludes the paper and gives scope for future works.

2 Literature Survey

Attendance can be recorded using numerous biometric approaches like retinal scan, fingerprint, iris recognition, face recognition, etc, that are used by many systems. D. Feng et. al.[1] worked on marking the student attendance, where the fingerprint is used. The major drawback of this system is the time taken to process and the need for users to physically go to the installed machines. M. Ali et. al. [2] used an iris-based attendance system. The issue with this approach is its sensitivity to environmental factors such as illumination changes.
S. Joardar et. al.[3] have developed an attendance management system on the basis of the palm dorsal subcutaneous vein pattern of individuals. PDSVP is used as a physiological biometric feature.
Jomon Joseph et. al.[4] the author proposed a face recognition attendance system using Eigenface recognition. However, this system is sensitive to faces in the background and head orientation and will fail if the person is wearing eyeglasses

or even facial hair etc.
Evta Indra et. al.[5] used Haar-like features to recognize faces of the students where they have to line up in front of the camera at a fixed position which is to be installed externally in the institute. But B.Tej Chinimilli et. al. [6], proposed a system with the state-of-the-art results using Haar Cascade and Local Binary Pattern Histogram (LBPH) algorithm.
Alghali et. al. [7] have presented a model that will identify individuals in every set of conditions especially in pose illumination conditions. A CNN architecture is applied to the complete system. The recognition operation is done by identifying indexes with a 99% accuracy.
In Jinhua Zeng et. al. [8], 33 face images were taken for forensic identification. It showed exceptional results on standard datasets such as labeled faces in the wild and YouTube datasets. The accuracy is similar to human prediction at 99%. The wide variability of face appearance and background noise of the image will be the challenge for face recognition[11].
In [12], the authors introduced a new facial recognition technique, that will compute 25 local autocorrelation coefficients. The proposed method in [13] includes detection and recognition for visually impaired people. In [14], the authors proposed an analysis of the Fast Gradient Sign Method. Used for biometric authentication. In [15], the authors proposed an open-source tool MaskTheFace to create a masked dataset. [16] proposed a modified CNN by adding two normalization operations to the layers. In [17], the authors proposed a face recognition algorithm The main aim of the system is to reduce the feature. In [18], the authors proposed an expression invariant 3D face recognition system. In [19], the authors proposed a model using PCA under the broad heading of factor analysis. In [17], the authors of the article propose a 3-D pose invariant approach based on subject-specific descriptors.

3 Methodology

The proposed method contains a face recognition framework integrated with the web application and the mobile android app, where the web application is built using the MVT (Model View Template) framework and a FaceNet[10] with InceptionResnetv2 framework in python is integrated. The Multi-Task Cascaded Convolutional Neural Network was used for face detection and a deep convolutional neural network for face recognition. Compared to the monotonic grayscale-based transformations, this detector is robust and performs well in terms of detection accuracy.

Figures 2 represents the user journey flow. Whenever the user accessing an application through the web app they will navigate to the home page (Which contains two features: Mark Attendance and Login) and when they are accessing the mobile app, an OTP is sent and verified, and then the user is redirected to the home page. The user employees have the privilege where they can also login into the app either with a face id or with login credentials to check their attendance log, request for leave, modify details/check the attendance log of

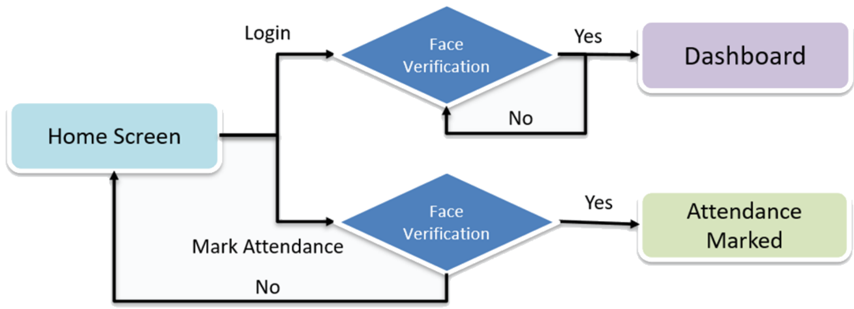

Fig. 2. Flow diagram of the user journey for the joint application

sub-level employees under their respective user access level, upcoming holidays, etc. The logged-in device location is tracked by the system. The user can mark attendance even without logging in and the admin(s) can monitor user activity.

The entire system's basic unit can be considered as the face verification/authentication unit. For the basic unit, a target face feature embedding is given and directly the recognized faces from the database are fetched.

Registering Employees: Each user can only be invited by a user with a higher access level. The invite is sent via email with an auto-generated username, password, registration link which redirects the user to the app where they can update their credentials and register their face in different angles. In order to feed the facial information into the backend database, an approach is employed where input is taken from the capturing module. In this module, a video camera is accessed in the front end interface and the individual frames from the real-time stream are captured and given to the processing module. This module performs: Face detection (identifying the placement of the face in the separated frame from the live stream and resizing the frame to fit the face only) from the frame using MTCNN (Multi-task Cascaded Convolutional Networks), a powerful face detector that has high detection accuracy. The faces stripped from the frame can be positioned in a way that the face verification unit might mistake it to be a different user. As a result, the picture used as input to the basic unit might be in a different state from the image to be analyzed. To resolve that issue, repositioning and alignment of the face detected in the image is done, which increases the accuracy of the face recognition model raising it from 98.9% to 99.7%. Here certain facial feature points that exist on the face are used to distinguish the face, its boundaries, the location of eyes, mouth and lips and then rotation is done on the face such that this transformation is similar to a perfectly centered face as possible. Therefore the face verification unit will not classify the projected faces as different. As a result, the redundant data that is not required for the face authentication unit is removed and the amount of pixels the model needs to work with is reduced, enhancing overall throughput. The processed image is taken and mapped into a compact Euclidean space. We are extracting the high-quality facial feature embeddings using a deep CNN, FaceNet model with an Inception-ResNetv2 framework and then represent the face as a 128-dimensional NumPy array that doesn't contain any redundant information of the face. The CNN's convolutional base is frozen which is responsible for the main feature extraction and the last fully connected layer is removed. The model used for face recognition is a deep convolutional network trained using a triplet loss function[10]. No intermediate layer is there to extract the feature set. These embeddings are stored in an array as a cube in the facial database table instead of storing images directly with the corresponding employee's id. Hence the required data of all the registering employees are fed to the system.

Marking Attendance: We employ motion detection when detecting the user's face in order to eliminate the possibility of proxy (holding a picture of another user's face) during authentication. The extracted face encodings of the user trying to mark attendance are taken from the live stream in a similar fashion as explained above and fed to the face authentication unit. The face authentication unit uses Euclidean distance similarity to find the distance between the vectors as a measure of facial similarity. Based on this facial similarity metric, identification of the face can be done. The distance similarity function is as below, using which, similar facial embeddings can be identified:

$$d(p, q) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}. \tag{1}$$

The basic unit takes the input encodings and compares them with the existing system database. The unit decides if they are the same person or different based on the threshold value which is set to 11.370 units. A similarity measure more than the threshold value is considered as a different face. The threshold value of '11.370' is fine-tuned with our custom dataset based on a decision tree algorithm as it splits the database where the gain is maximized for that particular value. The decision tree is built using a part of the database mentioned in the next section by forming a dataset with pairs of similar facial images as well as pairs of different facial images whose distribution is shown in the figure 3, where the blue line depicts the similar faces and the orange line depicts the dissimilar faces. As seen in the figure, below approximately 11 units, it can be considered of similar faces. Hence, we tuned the threshold to 11.37 (which is an output from the decision tree) as it gives the most optimised results without compromising the performance of the method as a whole.

To recognize the user, a query is used to directly fetch the employee id of the recognized face by contacting the backend database. This makes the overall process faster than that of the literature and also reduces the computational power required. Then automatically depending on the corresponding employee's id, the attendance is updated with the time stamp in the database's timesheet.

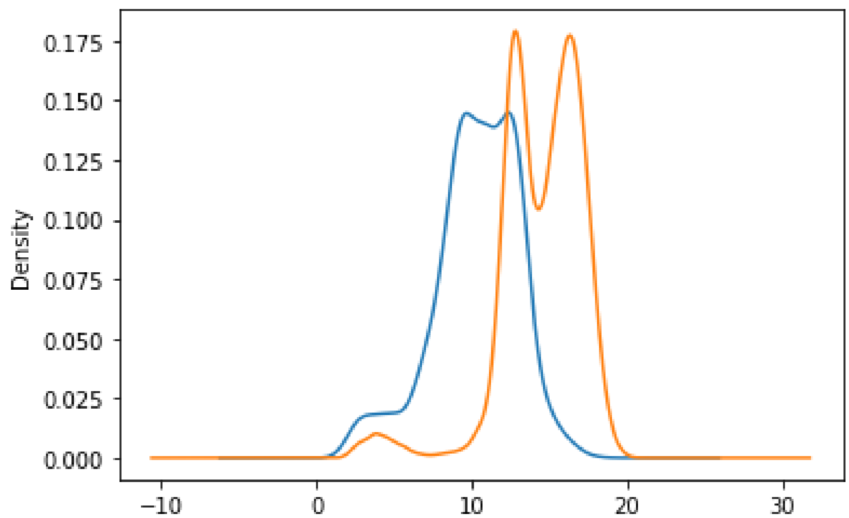

Fig. 3. Distribution plot of corresponding similar and dissimilar faces of the created dataset with Indian ethnicities.

Fig. 4. Output of capturing module

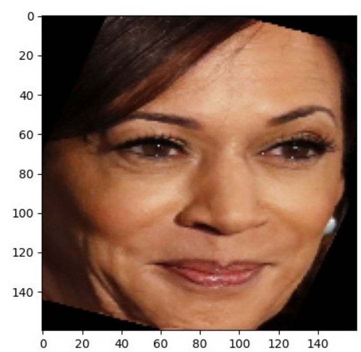

Fig. 5. Output of processing module

Fig. 6. Detected face Vs Registered face

4 Experimental Results and Discussion

In a real-time situation, gathering large volumes of data is a tough, monotonous, and time-consuming process. The initial samples collected were using various standard facial datasets and on the LFW (labeled faces in the wild) dataset (has 13,233 images), the developed method gave 99.7% accuracy. But there isn't any standard dataset that is best suited for this project since various angles, expressions, illumination, obstacles, all for a single individual are desired, however, most typical datasets contain millions of single faces, merely various illuminations, expressions of different humans. Illumination and pose of the input facial images affects the efficacy of face detection which affects the overall accuracy. Our pre-processing methods are designed to eliminate pose artifacts(through the face alignment and cropping part) but large variations such as looking sideways/occlusions due to objects/accessories not accounted for can affect the detection of faces.

The optimized and altered in the wild database was created from scratch which consists of 830 individuals with images in the wild ranging from 5 to 20 per individual(out of which one or two images is kept aside for validation of the method) in various expressions, lightings, with and without glasses, angles, projections have been taken for this experiment. In the database, 135 individuals are considered for the purpose of intruder testing, where the method's reliability of whether or not it recognizes the unknown person as its own is tested and hence not stored in the main database of the system. The majority of the database consists of Indian faces,which does not overlap with LFW. Our dataset, unlike other standardised benchmarks, does not solely concentrate on caucasian faces and by extension, helps our proposed method capture facial features of Indian ethnicities in greater detail.
The data is fed to the database and then the validation dataset is fed in the application layer to get the test accuracy and the evaluation metrics of this developed method. Testing of this method is done with the use of live video streaming where employees as well as the people who are foreign to the system use the app. The faces that are foreign to the system are detected and saved which helps to identify intruders and also reduce the incorrect authentication of the users to a Jane Doe (an intruder/faces foreign to the system).
The model and similarity measure used in this method gave more accuracy than those with other combinations (Example: FaceNet with InceptionResnetv2 and cosine similarity metric, OpenFace and euclidean distance, etc). The similarity measure threshold was finetuned in face recognition with the custom dataset with south asian faces in the wild. After experimenting with different face recognition models in combination with a detector, the best combination that gives an accurate performance is used, which is MTCNN for face detection and the FaceNet with InceptionResnetv2 model for face recognition. MTCNN outperforms haar cascade, dlib based approaches in the proposed architecture. This detector also detects the user from a distance upto 16 feet given the most of the face is covered in the stream, head movements and extracts facial landmarks like

eyes, nose, mouth which is important in the processing module for face alignment as that improves the overall accuracy, precision of the facial recognition module. Since it performs well compared to those of the other modules. Even though the method gives a state of art accuracy, the entire system architecture results in an improved true positive, false negative, false-positive rates for both existing registered databases and the foreign individuals who are unknown to the system.

Results: Table 1gives the experimental results as compared to that of the existing literature. The existing architecture results can be found in [6] for true positive, false positive rates. Table 3 gives the results obtained from the developed method which is tested from the 830 individuals of the dataset.

Table 1. Benchmark Comparison

Performance Metric	Existing Method	Developed Method
True Positive Rate	77%	96.14%
False Positive Rate	28%	14.61%
False Positive Rate(User foreign to system)	14%	6.66%
False Negative Rate	23%	3.86%
True Negative Rate	72%	85.39%

Table 2. Confusion Matrix

	Predicted: No	Predicted: Yes
Actual: No	TN: 222	FP: 38
Actual: Yes	FN: 22	TP: 548

The performance of such attendance marking systems is based on how accurately the users are authenticated and they are measured using the true positive rate, true negative rate, false positive rate, and false negative rates. The true positivity rate of the proposed method is 96.14% (19% greater than that found in existing literature). The tests conducted accounted for varying environmental setups ranging from illumination changes(Outdoors, Indoors, partially outdoors) to presence of occluding accessories, glasses, cap, with dyed hair etc. Even cases where the user may attempt to proxy another user's face with a picture were added in order to raise the robustness of the method. For faces that are foreign to the system, the false positivity rate is 6.66% which is 8% less than that of the state of art, making the system more secure. The false positive rate for the users of the system is 14.61% while it is 28% for some of the existing systems. The false-negativity rate was found to be 3.86%, making the system architecture more reliable as the registered employees will not be rejected wrongfully. The true negative rate is 85.39%, which implies the method will detect the non-users of the system correctly. The overall system time complexity is also improved compared to the literature as well the recognition rate and its instant updation to the main database takes approximately 5 seconds, unlike an existing attendance system[9] which takes 120sec.

Table 3. Experimental Results with finetuning

Framework	Accuracy
Facenet	92.9%
Facenet Fine-Tuned	99.7%
VGG-Face	95.5%
VGG-Face Fine-Tuned	99.2%
OpenFace	93.8%
OpenFace Fine-tuned	98.7%

The method was implemented using various state-of-the-art frameworks for face recognition like Facenet, VGG-Face, OpenFace, and their fine-tuned version with respect to threshold value with the custom dataset in the wild. The basic authentication unit was fine-tuned for optimal threshold value with respect to the custom dataset. Their corresponding accuracies as shown in table 3. The state-of-the-art frameworks were pre-trained with mostly caucassian faces and hence gave less accuracy when tested with Indian faces due to difference in detail of the facial features.

The developed method is working with real-time input and obtaining robust and accurate results with less time and space complexity without requiring a larger database. The existing methods that use face recognition need a large amount of data per face which is 70 images per person which thereby increases the false positives inside the system whereas the developed method requires fewer facial images per person and yet gives a more optimal false positive rate within the system than literature.

The proposed efficient architecture also reduces the risk of a data breach and mitigates a wide range of security issues. It has an improved face recognition runtime than existing methods where even if a separate microprocessor is used which takes time to load, the weights of the model and process the recognition module. Here, the recognized individual id is directly fetched from the database. The proposed architecture, despite using a Deep CNN requires less computational power, is less time consuming for the user and is capable of handling a large data without compromising its performance, making it suitable for use in a large organization, unlike the existing methods in the literature considering the number of students, employees, teaching and non-teaching faculty, and departments.

5 Conclusion

This article describes a smart end-to-end approach for an automated method that successfully marks attendance of the staff, maintains a log of details of the staff, gives various features for the staff to check upcoming holidays, apply for leave, attendance log inquiry etc. The method uses an MTCNN, InceptionResnetv2 based Facenet framework integrated with an MVT framework architecture on a web-based as well as an android mobile platform. It is a multi-purpose application for the employees in an institute/organization/university/office that makes most of the processes done online without any physical contact and without intermedial delay or any human intervention. It also follows a specifically designed user access management policy. It is easy to deploy and maintain and in order to deploy the system in the institute or workplace environment, no specific hardware gear is required. The end product is a user-convenient system that makes a contribution to the creation of a better, healthier educational institute/work environment. Furthermore, the developed method can be used in a variety of fields where automatic attendance management is required or made more efficient without physical intervention. The users won't need to worry about their privacy since the original face images aren't saved in the system server/backend database. External hardware gear isn't required to be installed or to come in contact with. The attendance records for all dates are likewise stored in the system. As a result, it will be easy to keep track of attendance data.
In the future, for augmenting the core algorithmic performance of the system architecture, the current network used in the method can be retrained over an additional structured dataset of the current version and experiment it with more benchmarks. In the processing module, detection and alignment of faces in the wild such as conditions of strong sunlight and other illumination challenges can also be improved. Also, the scope of this method can be expanded by enabling it to handle multiple users within a single frame for effective use in public locations.

References

1. D. Feng, P. Wang, and L. Zu: Design of Attendance Checking Management System for College Classroom Students Based on Fingerprint Recognition. In: Chinese Control And Decision Conference (CCDC), pp.555-559 (2020)
2. M. Ali, H. Usman Zahoor, A. Ali and M. Ali Qureshi: Smart multiple attendance system through single image. In: IEEE 23rd International Multitopic Conference (INMIC), pp. 1-5, IEEE (2020).
3. S.Joardar, A.Chatterjee, and A.Rakshit: A real-time palm dorsa subcutaneous vein pattern recognition system using collaborative representation-based classification. In: IEEE Transactions on Instrumentation and Measurement, vol. 64, no. 4, pp. 959–966, IEEE (2015)
4. Jomon Joseph, K.P.Zacharia: Automatic Attendance Management System Using Face Recognition. In: International Journal of Science and Research, IJSR, vol. 2, issue 11, pp. 327 – 330 (2013)

5. Evta Indra, Muhammad Yasir, Andrian, Delima Sitanggang, Oloan Sihombing, Saut Parsoran Tamba, Elviana Sagala: Design and Implementation of Student Attendance System Based on Face Recognition by Haar-Like Features Methods. In: 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT), pp.336–342 (2020).
6. Bharath Tej Chinimilli, Anjali T, Akhil Kotturi: Face Recognition based Attendance System using Haar Cascade and Local Binary Pattern Histogram Algorithm. In: Fourth International Conference on Trends in Electronics and Informatics (ICOEI)(48184), 2020, pp. 701-704 (2020).
7. Mohammed Alghaili, Zhiyong Li, and Hamdi A.R.Ali: FaceFilter: Face Identification with Deep Learning and Filter Algorithm, Hindawi Scientific Programming Volume 2020, Article ID 7846264, 9 pages (2020)
8. Jinhua Zeng, Xiulian Qiu, Shaopei Shi: Image processing effects on the deep face recognition system. Mathematical Biosciences and Engineering, vol. 18, issue 2, pp. 1187–1200. (2021)
9. S.Dev and T.Patnaik: Student Attendance System using Face Recognition. In: International Conference on Smart Electronics and Communication (ICOSEC), pp. 90-96, IEEE(2020)
10. F.Schroff, D.Kalenichenko and J.Philbin: FaceNet: A unified embedding for face recognition and clustering. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 815-823, IEEE (2015)
11. M.Nandini, P.Bhargavi, G.Raja Sekhar: Face Recognition Using Neural Networks. In: International Journal of Scientific and Research Publications, vol. 3, issue 3, (2013)
12. F. Goudail, E. Lange, T. Iwamoto, K. Kyuma and N. Otsu: Face recognition system using local autocorrelations and multiscale integration. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, issue 10, pp. 1024-1028, (1996)
13. Mr.G.Arun Francis, Dr.P.Karthigaikumar and Mr.G.Arun Kumar: Face recognition system for visually impaired people. In: Journal of critical reviews, vol. 7, issue 17 (2020)
14. Arbena Musa, Kamer Vishi and Blerim Rexha: Attack Analysis of Face Recognition Authentication Systems Using Fast Gradient Sign Method. In: Applied Artificial Intelligence (2021)
15. Anwar, Aqeel, Raychowdhury and Arijit: Masked Face Recognition for Secure Authentication. arXiv:2008.11104 (2020)
16. M. Coşkun, A. Uçar, Ö. Yildirim and Y. Demir: Face recognition based on convolutional neural network. In: 2017 International Conference on Modern Electrical and Energy Systems (MEES), 2017, pp. 376-379 (2017)
17. Hongling Chen and Chen Haoyu: Face Recognition Algorithm Based on VGG Network Model and SVM. In: 2019 3rd International Conference on Machine Vision and Information Technology (CMVIT 2019) Guangzhou, China, vol 1229 , pp. 22–24 (2019)
18. A.M. Brostein,M. M. Brostein and Ron Kimmel: Three-Dimensional Face Recognition. In: International Journal of Computer Vision, vol. 64, issue 1, pp. 5–30 (2005)
19. M. Khan, S. Chakraborty, R. Astya and S. Khepra: Face Detection and Recognition Using OpenCV. In: 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), pp. 116-119 (2019)
20. Naeem Iqbal Ratyal, Imtiaz Ahmad Taj, Muhammad Sajid, Nouman Ali, Anzar Mahmood and Sohail Razzaq: Three-dimensional face recognition using variance-based registration and subject-specific descriptors. In: International Journal of Advanced Robotic Systems, vol. 6, issue 3 (2019)