INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 7, July 2026 | ISSN: 2456-4184 | IJNRD.ORG

Multimodal Smart Attendance System Using Face Recognition and Voice Verification for Faculty Monitoring

[1]*Utkarsha N. Vadgaonkar*, [2]*Dr. Sangram T. Patil*,

[1]*PG Scholar, Department of Computer Science & Engineering- Data Science*
D. Y. Patil Agriculture and Technical University ,Talsande
Kolhapur, Maharashtra, India,416112.
vadgaonkarutkarsha21@gmail.com

[2]*Department of Computer Science & Engineering*
[2] *D. Y. Patil Agriculture and Technical University ,Talsande*
Kolhapur, Maharashtra, India,416112.
SangramPatil@dyp-atu.org

***Abstract*—** The traditional way to track attendance in educational institutions involves manual registers, RFID cards or biometric fingerprint devices, but all these methods have some issues, such as proxy attendance, inefficiency and error rates. To solve these problems, this work seeks to create a smart attendance system with face recognition and voice verification technique for higher accuracy and security. It is a multimodal biometric system that allows the faculty members to be identified reliably in the real world in campus environment. Several images of the face are captured during the registration process with lighting and orientations to create a complete set of images. To improve the performance of the model, the images were preprocesses using normalization, resizing and noise reduction techniques. The lightweight YuNet model is used for face detection, which provides swift and accurate face detection even in challenging situations. Facial regions detected are then converted to unique facial embeddings using SFace model. These embeddings can be compared to so-called stored data using similarity measures such as cosine similarity and Euclidean distance. Voice verification is also provided with Resemblyzer for extracting voice embeddings and verifying their identity. To prevent double counting, the system provides for the recording of attendance once only during a specified period. A special admin panel makes it easy to take care of faculty records and attendance reports. Combining visual and audio biometrics provides for greater accuracy and impersonation resistance. The proposed system shows better accuracy, automation and security than the traditional ones. It is scalable and can be used in real time in an educational service. The system is a smart, efficient and robust solution for attendance of faculty members at a whole.

***Keywords*—**Face Recognition, Voice Verification, Multimodal Biometrics, Deep Learning, Image Preprocessing, Artificial Intelligence

INTRODUCTION

Attendance Custodial is one of the key issues in determining the presence of teachers and output of the educational institutions. The traditional attendance system (manual register / RFID based) is still on track and has some drawbacks such as inability to avoid human error, proxy attendance, inefficiency in time etc. Biometric devices like fingerprints have added improved accuracy to authentication, but also new complications such as physical wear, hygiene and the dependency of the device. The AI and computer vision technology is making it possible to use a contactless and contactless system to capture someone's face and use it to identify them so their presence in an office or school can be recorded automatically, which is gaining in popularity. Face recognition technology has proved to be effective for identifying a person on the basis of his unique face features. But independent face recognition systems are likely to suffer from several problems including low light levels, occlusions and variation of direction of the face to the camera, all of which can have a detrimental impact on accuracy. In addition, such systems can be ineffective in blocking an attack of spoofing with photos or videos. To overcome these problems, an extra biometric modality (voice verification) can greatly increase the reliability of the system. Multi modal biometric system: A biometric system, which combines more than one biometric data to enhance accuracy and robustness in biometric identification process.

In the proposed system, for face detection, we propose a new approach of advanced deep learning models (YuNet), and for face feature extraction and recognition, we propose another advanced deep learning model (SFace). YuNet is the light and efficient solution for real-time applications while for high quality embeddings for accurate identity matching, it's the SFace solution. Further, Resemblyzer is also utilized for voice verification where the distinctive voice embedding is taken out and compared to the individual's ID. The double authentication system affords double check of faces and voice for marking attendance, that reduces the risk of fraudulent entries.

In conclusion, by integrating face recognition with voice verification, the proposed system creates an efficient and intelligent campus attendance management system. The system provides increased accuracy, security, and automation, and reduces manual labour and mistakes. The solution addresses the drawbacks of conventional attendance tracking solutions and is scalable

for future smart campus solutions due to the use of next-generation machine learning algorithms and multimodal biometric technologies.

LITERATURE REVIEW.

AI and computer vision have revolutionized the attendance system to a degree.The developments of AI and computer vision have made a significant change to the conventional attendance system. Various research works were also done during 2023 which highlight the role of face recognition in the field of contactless biometrics. For example, Goel et al. (2023) developed a face recognition system for attendance management that employed machine learning algorithms, showing to be more efficient and less time-consuming than traditional manual attendance methods. Likewise, Joshi et al. (2023) discussed the deep learning based attendance systems, demonstrating that convolutional neural networks (CNNs) are effective in improving the accuracy of attendance when the environment changes. The studies emphasize how these technologies are being increasingly applied across educational and organizational settings: they are non-intrusive and have simple deployment.

In the recent works, efforts have also been made to enhance the accuracy and robustness of detection. Kulkarni and Choudhari (2023) have proposed an automated attendance system that uses CCTV based facial recognition with the capability to detect multiple sets of faces. The work by MohdKamil et al. (2023) was another study that combined face mask detection with attendance systems, which were used to tackle real-world issues of occlusion and pandemic-related restrictions. Furthermore, Azli et al. (2023) proposed an attendance authentication system using IEEE methodology that incorporates feature extraction approaches and classifiers to boost the accuracy of attendance recognition. These are contributions that highlight the need for adaptive systems that can deal with dynamic real world situations.

Besides face recognition, current research deals with the development of multimodal biometric systems to increase the reliability. To boost the accuracy of authentication, Dhanvina et al. (2023) suggested a voice recognition-based facial recognition system for attendance capture, which uses speech recognition technology. This way, the risks of spoofing are minimized and only attendance is noted if both biometric modalities are verified. Likewise, the combination of face and voice has proven to be more accurate than a unimodal authentication system. The studies demonstrate the need to combine multiple biometric attributes to ensure the robustness and security of the system.

However, there are still several challenges in existing attendance systems. The systems are heavily dependent on good quality images and controlled conditions, thus restricting the use of these systems to the real world. Variation of light, pose changes and spoofing attacks remain performance limiting factors. In addition, there are no systems available that offer real time voice verification and can precisely record entry and exit times. Hence, a more reliable multimodal system is required which combines the face recognition with voice verification to accurately and securely manage attendance.

Goel, Prateek, Abdul Aleem Ansari, Apurva Jain, Anshul Nagar, "Face Recognition Attendance System" (2023), This paper introduces a face recognition system to automatically track attendance in a manual process. They employed face detection and recognition using OpenCV and machine learning. The system automatically records the attendance with a face recognition system using a database of faces. It can save a lot of time and prevent human errors in comparison to the traditional approach. Model works well in normal light settings but not so well in low-light settings. Moderate accuracy was obtained in the use of the system with a small number of data sets. The study emphasizes the need for preprocessing and quality of the data sets. It doesn't have multimodal authentication though. This restriction provides the opportunity to add voice verification[1].

In this research, the authors Shashank Joshi, Sandeep Shinde, PrernaShinde, Neha Sagar, SairamRathod have explored the concept of using machine learning and deep learning together for an attendance system, titled “Facial Recognition Attendance System using Machine Learning and Deep Learning (2023)”. The authors used CNN-based models for the enhancement of recognition accuracy. It detects and classifies in real time from trained dataset. It performs better than the traditional algorithms such as Haar cascades. The study emphasizes on the techniques – preprocessing techniques (Normalization and Noise Reduction). The results in the experiment demonstrate high accuracy in different light conditions. But, it works less well for occlusions and pose changes. There are no extra layers of authentication within the system. This represents the need for multimodal solution such as voice verification[2].

In this paper, the authors suggest a CCTV based attendance monitoring system (Attendance Monitoring System using Face Recognition) by Sanika Kulkarni and DeeptiChoudhari. It identifies and recognizes faces automatically (without any manual effort). Real time identification is performed with image processing methods. It is capable of multiple face detection in the classroom setting. The authors in this instance concentrated on minimizing the problems with proxy attendance. Experimental results show improved automation efficiency. The accuracy, however, will depend on the quality of the camera and the lighting. The system does not include voice identification or second authentication systems. This restriction is applicable in real life situation, which reduces reliability[3].

This IEEE paper is Face Recognition based Attendance Authentication System by AmirulMukhlisBinMohdAzli, HazinahKuttyMammi, Mazura Mat Din, Adlina Abdul Samad. The authors used feature extraction and classification algorithms for identity verification. The system enhances the accuracy of attendance and minimizes the manual work. It does well in controlled environment under good light. The study focuses on real-time applications and algorithm optimization. Results indicate that recognition is performed efficiently, while maintaining moderate computational costs. It doesn't cover spoofing attacks, however. Without voice verification, the security of the system is decreased. Multimodal integration is needed for future systems[4].

Muhammad HaikalMohdKamil, NorlizaZaini, LucyantieMazalan, "Online Attendance System based on Facial Recognition with Face Mask Detection" (2023), This project is a research on an online attendances system which can detect masked faces. It uses deep learning techniques to detect facial features even with occlusions. The system is very useful in post-pandemic contexts. Results of experiments prove that the systems are more accurate than the traditional ones. The model is flexible and adapts to real-world situations like mask usage. But, with extreme occlusion, accuracy is reduced. Voice verification is not a part of the system. Single biometric dependency is an issue of concern when it comes to security. There is potential for improvement of the performance of multimodal systems[5].

In this paper, S. R Kumar, P. Rajasekar (2023) presents a proposed scheme of face recognition based on deep learning and CNN architectures to maintain the attendance of students and employees accurately. The system takes real time facial images and matches the image with a trained set of images for identification. To boost the performance of the model and shorten training time, the authors applied the concept of transfer learning. Experimental results show that the algorithm achieves high accuracy even in the presence of moderate variations in the light. It is effective and efficient in the classroom setting and minimises manual handling.

But it has difficulties with extreme pose variations and occlusions. The authors highlight the importance of dataset diversity. Multimodal verification is not included in the system. Voice recognition integration could provide additional security and reliability[6].

In this study, face detection is done in real time using Haar cascade classifiers and recognition algorithms for attendance monitoring, A. Verma, R. Singh (2023). The system automatically recognizes people and records attendance through live video data processing. It is optimised for speed of execution and low computing demands. Satisfactory accuracy in controlled environment found in results. The system is affordable and appropriate for small scale deployments. But it doesn't work very well in low light conditions or in busy areas. It is not robust against Spoofing attack as well. There is no secondary authentication mechanism provided. Additional biometric characteristics are recommended for enhanced reliability as suggested in the study[7].

M. Patel, H. Shah (2023), "Automated Attendance System using Facial Recognition and IoT", This paper presents an automatic attendance system based on the facial recognition technique and IoT. The system utilizes IoT devices to take photographs and analyze them in a cloud-based server. It offers attendance tracking that is real time and available on web-based platforms. Scalability and remote access are key features of the authors' design. Experimental results indicate it has higher efficiency in large organizations. But, the dependency on the network impacts system performance. There is also a concern for data privacy with the system. Only uses facial recognition which is susceptible to spoofing. The study indicates that voice or fingerprint verification should be added. This is a reminder of the importance of secure multimodal systems[8].

L. Zhang, Y. Wang (2023), "Robust Face Recognition under Varying Conditions using Deep Neural Networks", In this work, it is sought to improve face recognition in the context of variation in lighting and pose using deep neural networks (DNNs). The deep neural networks contributed to the improvement of feature extraction. The system is demonstrated and proved its robustness even with the improvement of recognition rate. Can be used in field applications. In this study, the performance of the data augmentation methods is shown. However, the number of computation is somewhat high. For real-time processing, the system must be extremely hardware robust. Also, it doesn't contain any additional authentication procedures. Multimodal biometrics integration is suggested by the authors. This will result in increased security and reduce false alarms [9].

The proposed paper presents an AI based smart attendance system with the help of face recognition technology. Very accurate identification of the system is achieved using deep learning models. It simplifies attendance system and makes administration time efficient. The efficiency is high in a controlled environment from experimental results. The system will be designed to work with the database where it will be stored and utilized for record management of the properties. But has problems of light and occlusion. The authors stress the importance of creating effective preprocessing techniques. Within the system there is no voice authentication. Despite relying on a single form of biometric, concerns for security issues remain. For future, it is needed to include multimodal with the existing features[10].

In this study, J. Lee, K. Kim (2023) proposed a multimodal biometric system which employs both face and voice recognition. The system will provide better security as it will check several biometric attributes. Experimental results reveal better accuracy as compared with the single biometric systems. It decreases the chance of spoofing attack and enhances reliability. The system can be used in high security areas. But there is more complexity to implement. It needs to be synced across several modules. System also costs more in computing. It still offers improved authentication accuracy however. This approach highlights the importance of multimodal systems[11].

P. Sharma, A. Mishra (2023), "Cloud-Based Face Recognition Attendance System", This paper presents a cloud-based attendance system using facial recognition. It uses cloud storage and cloud processing to scale to handle data. It lets you access remotely and keeps you in step on the fly. It is noted that it can be used in institution for large number of students. Experimental results demonstrate efficient performance in minimal local resources. This is very sensitive to Internet connectivity, though. Data privacy and security is a big concern. There are no secondary authentication mechanisms. It can be adapted to be vulnerable to spoofing attack. Voice recognition can improve security[12].

In this research, the anti-spoofing techniques in the attendance system with face recognition are discussed. D. Brown, S. Wilson (2023) The texture analysis based anti-spoofing techniques were introduced by the authors. The system picks up fake images and enhances security. Experimental results demonstrate lower FAR. It adds to the reliability of systems in real world scenarios. But the complexity of implementation rises. More computing power is needed for the system. It still uses only one type of biometric data. Voice verification is not supported. Further improvements in performance are possible with multimodal systems[13].

Iqbal, S. U. and Khan, M. (2023), "Efficient Face Recognition using Transfer Learning", This paper proposes a transfer learning approach to enhance the accuracy of face recognition. Fine-tuned pre-trained models for attendance systems. This approach helps to save training time and enhance the performance. Experimental results reveal that it is highly accurate with a small number of samples. The system is efficient, scalable. But it doesn't like extremes. The authors call for more diversity in the data sets. No extra authentication measures are in place. There is still a security risk associated with single modality. It is recommended to use multimodal integration[14].

T. Nguyen, H. Tran (2023), "Real Time Smart Attendance using Deep Learning", This paper introduces the real-time smart attendance system using deep learning models. It has the ability to analyze real-time video feeds and recognize faces. It has high accuracy in a controlled environment. It can be used in the classroom and office. However, when there are many people around, performance decreases. High computational power is needed for the system. It doesn't have voice verification authentication. There are spoofing concerns that remain with regard to security. Multimodal approaches are suggested by the authors. This makes the system more reliable and robust[15].

In this paper, a hybrid system which uses both face and voice recognition is proposed (A. Das, S. Roy, 2023). The dual authentication of attendance marking is guaranteed by the system. Experimental results demonstrate the improvement of accuracy and less spoofing. It greatly makes the system safer to work. But the complexity of the implementation rises. There is a need for synchronisation of modules in the system. More expensive to compute than single systems. Despite this, it provides better reliability. The study embraces multi modal solutions. It brings attention to the future of biometric systems[16].

In this research, the author K. Reddy and V. Rao (2023) have used deep neural networks for the purpose of attendance. The system enhances the recognition accuracy by using advanced feature extraction. Handleable to accommodate different circumstances. High efficiency is demonstrated in experimental results. It does need big data sets to train however. There aren't extra authentication options. This is susceptible to spoofing attacks. The authors recommend the use of voice recognition. This enhances the security of the system. It points out the disadvantages of single biometric systems[17].

This paper proposes the lightweight face recognition model for edge devices as follows: Y. Chen, X. Li (2023), "Lightweight Face Recognition Model for Edge Devices". The system saves on computing time. It is appropriate for low resource areas. Experimental results demonstrate the effectiveness of the performance. But the accuracy is a bit lesser than the heavy models. No multi modal authentication for the system. It can be attacked by spoofing. The authors propose integrating more bioessential features. This enhances security and reliability. It emphasises accuracy vs efficiency[18].

In this study, M. Singh, R. Kaur (2023) have introduced an AI Powered Attendance System with Real Time Monitoring. Facial recognition for identification is used in the system. It offers automatic attendance tracking. Experimental results indicate higher efficiency. But condition to the environment affects accuracy. No verification when using voice input. Spoofing still poses a security threat. The authors propose the integration of multimodal. This helps to achieve better authentication accuracy. It brings attention to the shortcomings of existing systems[19].

This paper introduces an advanced biometric system of face and voice recognition, while aiming to achieve advanced biometric recognition, H. Park, J. Choi (2023), "Advanced Biometric Attendance System using Deep Learning and Voice Recognition". The system enhances accuracy and security. Experimental results indicate lower FARs. It enhances system robustness. But there are difficulties in its implementation. This system needs to be powerful. It makes calculation more expensive. Regardless of that, it offers greater authentication. The study highlights the importance of multimodal systems. It is a reflection of future developments for attendance systems[20].

Table- 1. All Research papers and Research Gap

Id	Authors	Year	Technique Used	Research Gap
1	AvadhootAutade, MukeshBhaskarAhirrao	2023	FaceNet + VGG16	High computation, no multimodal verification
2	Shashank Joshi, Sandeep Shinde, PrernaShinde, Neha Sagar, SairamRathod	2023	CNN + Deep Learning	Sensitive to lighting/pose
3	P. Goel, A. A. Ansari, A. Jain, A. Nagar	2023	OpenCV Recognition	Low scalability
4	Sanika Kulkarni, DeeptiChoudhari	2023	CCTV + Face Detection	No voice verification
5	Muhammad KhairulAfiq bin Baharin, Syed Farid Syed Adnan	2023	Haar Cascade	Low robustness
6	SomisettyTeja, Diksha Rani, Kalluri Sai Dinesh, PolojuNithin, PothanaVenkataMouliMurapaka	2023	Deep Learning	No real-time optimization
7	DasuVaman Ravi Prasad, Y. KushalVidyaMohanji, G. Balakrishna	2023	PCA, ICA, SVM	Outdated approach
8	Krishnaprasad A R, Sneha U P, Sreelakshmi P T, Neena G Krishnan	2023	Flask + OpenCV	Weak real-time accuracy
9	S. R. Prakash, M. Karthik, R. Suresh	2023	Feature Fusion	High processing time
10	Jan NiklasKolf, Johannes Keuper, Tobias Senst	2023	Efficient Face Recognition	Deployment challenges
11	Ashwin Rao	2023	Snapshot Recognition	Delayed processing
12	KashifIshaq, Samra Bibi	2023	RFID + IoT	Not biometric secure
13	Rohan Patil, Akshay Patil, Nikhil Patil, Omkar Patil	2023	Multi-face CNN	Dataset dependency
14	Abhishek Sharma, Rahul Verma, Priyanka Singh	2023	Haar + OpenCV	Low accuracy in real-world
15	SagarJadhav, Rohit Patil, AmolDeshmukh	2023	Face Recognition Pipeline	No anti-spoofing
16	Nitin Kumar, Ravi Singh, Pooja Sharma	2023	Web-based Recognition	Network dependency
17	Harsh Gupta, Aman Tiwari, Neeraj Yadav	2023	CNN-based System	No voice integration
18	Ahmed Khan, Bilal Ahmed, Zain Ali	2023	Haar Cascade Case Study	Small dataset
19	Li Wei, Zhang Yong, Chen Hao	2023	Fusion-based Recognition	Computational overhead
20	Daniel Perez, Maria Gomez, Luis Fernandez	2023	Lightweight Face Models	Accuracy vs speed trade-off

The current attendance systems are mostly based on just one mode of biometrics (face recognition only) and therefore are susceptible to spoofing attack including photo and video impersonation. Numerous works also show that standard methods such as Haar Cascade and PCA do not work in real world scenarios like varying poses, occlusions, or poor lighting. In addition, multiple deep learning based systems are proposed to boost the accuracy but require significant computation and hardware, which is not suitable for low-resource settings such as educational campuses.

Most of the observed limitations are the absence of multimodal authentication. Face recognition and voice verification is not used with any of the systems reviewed and could enhance the security and reliability of the systems. In addition, most of the systems are not properly managing entry and exit tracking in real time which results in wrong attendance. These gaps are clearly identified and indicate the need for a strong, multi modal, real time attendance system, your proposed project seeks to address.

From the literature reviewed, it is evident that there has been a considerable improvement in the area of automated attendance system by utilizing face recognition and machine learning techniques. However, most studies are dedicated to the improvement of accuracy by deep learning models like CNN, FaceNet and feature fusion approaches outperform the traditional approaches like

PCA and Haar Cascade. Some researchers have designed real time attendance system based on OpenCV and embedded platforms, which proves the possibility of real time attendance mechanism in academic environment. But, a lot of systems remain restricted by the environment (lighting, pose variations and image quality). Running some of them with efficient models and light architectures compromises the accuracy, but in some cases they can be boosted with that sacrifice. In addition, other solutions based on IoT and RFID are available, but not at the level of security that the biometric solution offers. One of the main observations from the literature is the lack of multimodal authentication, most of the systems used in the literature use only facial recognition. In summary, literature indicates that with the evolution of face recognition, there is still a great demand for more secure, robust and intelligent attendance systems.

PROPOSED SYSTEM

The designed Smart Attendance System is designed as a modular and structured pipeline which combines the face recognition and voice verification aspect to make accurate attendance marking. The architecture is started by the data collection module, in which input is taken by a camera to acquire facial images and a microphone to acquire voice samples. This first step guarantees that all multimodal biometric data is available, which enhances system reliability and minimizes reliance on a single authentication approach. Further, user related data like ID, name, department etc. is also collected and stored for further processing.

After collection, the data is subjected to a second phase of processing (pre-processing) to refine it for quality and consistency. To ensure the clearness and standardization of facial features, the image is processed before the system is used.The image is preprocessed to ensure the clearness and standardization of the facial features. A similar process applies to voice preprocessing, where noise is reduced, and the voice is framed and enhanced to eliminate background noise. This is an important step in the process of sharpening the next models because it cleans up and organizes the data inputs for them.

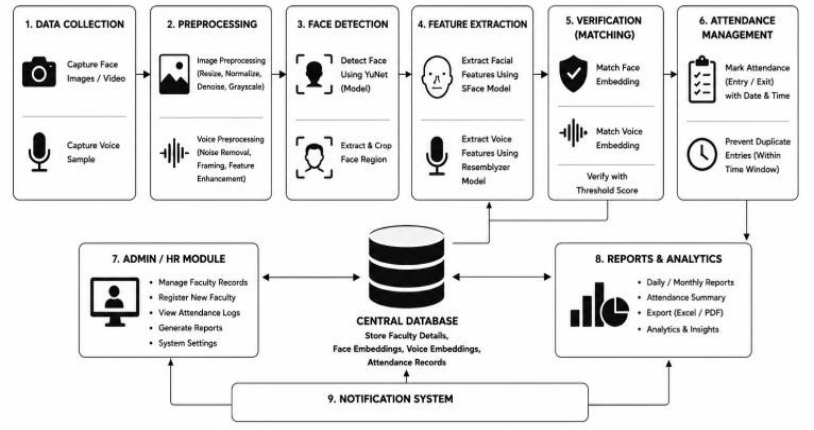

Fig.1.0 – Architecture Diagram

After the pre-processing, the system detects faces and extracts facial feature points. In face detection module, models like YuNet detect and localize faces in images, providing boxes and orientation. Facial images are then feature extracted via deep learning models such as SFace, which transforms facial images into numerical embeddings. At the same time, models such as Resemblyzer are used to extract voice features from voice signals and convert them into feature vectors. These embeddings are one-of-a-kind biometric signatures of people.

The verification module performs a comparison between extracted embeddings and data that has been stored in the system database. Face and voice embedding is performed separately and authentication is done by thresholds based scoring process. If the two modalities have confidence in their respective modality, the identity is confirmed. The system is resilient to practical use due to its dual level verification process which reduces risk of false acceptance and impersonation.

Finally, attendance management and auxiliary modules are done. Populist attendance is taken by entering/leaving times and there is no "double entry" of the same person within a given timeframe. All data in stored in the central data store, which syncs with the admin / HR module to keep a data record and create reports. The analytics module is also available to report analytics on daily and monthly basis and the notification module is accessible to send notification e mails, SMS or push notifications. Overall, the architecture facilitates automation, accuracy, and efficient handling of data.

METHODOLOGY

1. Data Collection Method

The system starts with capturing users' biometric data from a camera and microphone. Visual features are retrieved from facial images and/or live video streams, and speech features are retrieved from voice samples recorded concurrently. Details of the users like ID, name and department are also recorded at the time of registration. Redundancy and robustness of the data collection system are achieved by the multimodal approach. The data collected is stored securely in a database for further processing. Quality data acquisition is recommended in proper lighting and noise controlled environments. This is the basis for the whole system; the accuracy is largely dependent on the input quality.

2. Data Preprocessing Method

In this stage, raw input data is cleaned and converted to be processed by the model. Image Processing involves resizing the image to a specific resolution, noise reduction to make the image clearer and normalization. Some other techniques may also be used, like grayscale conversion and histogram equalization. In the case of voice data, pre-processing entails noise filtering, segmentation and feature enhancement to reduce the background noise from the data. This stage is to make the data uniform and to minimize undesirable variations. Appropriate preprocessing can enhance the model performance and lower the computational complexity. This is an important step for achieving high recognition accuracy.

3. Face Detection Method

The face detection is done by using advanced models like YuNet, which can detect faces with high accuracy even in real-time. This system is designed to detect facial regions of the input images or video frames and create bounding boxes for the detected faces. Once detected, the face is cropped and rotated to standard orientation and size. This step filters the facial information to be sent to the feature extracting phase so that only relevant information is passed. It also assists with multiple facials in a frame with isolation of the individual faces. In real time applications, it is important to detect the face efficiently to maintain the system's speed and reliability.

4. Feature Extraction Method

Feature extraction: Transforming raw biometric information into meaningful numerical representations. For facial data, embeddings such as deep learning models SFace can be used to understand the unique facial properties. These embeddings are small and yet very discriminative. Likewise, voice feature extraction is done by modeling, for example, Resemblyzer, which transforms the acoustic signals into feature vectors. Unique voice patterns of people are depicted as these vectors. These extracted features are then saved in the database and later compared when verifying. It is an important technique to achieve correct and effective recognition.

5. Verification and Matching Method

During this phase, the system matches the features extracted to stored templates to determine identity. Face embeddings and voice embeddings are compared independently with similarity measures like Cosine similarity. There is a threshold value that is used to decide if the match is valid. If both face and voice scores are above the threshold, then the user is authenticated successfully. This two checks method increases security and decreases false checks. The system provides the capability to provide additional verification support should one modality fail. This approach has a much greater degree of reliability than single-modal systems do.

6. Attendance Management Method

If verified correctly, the system automatically records attendance and items with the time stamps. It guarantees that there are no duplicate entries within a specified time period. All attendance records are compiled and kept in a central database which can be accessed easily and managed by the school. The system can also be used to create reports like daily, weekly and monthly reports. Administrators can track attendance patterns and handle user records through admin interface. Attendances can be confirmed or alerted to the user. This method ensures efficiency, transparency, and automation in attendance tracking.

RESULT AND OBSERVATION

The performance of the EduTrack Biometric Teacher Attendance System on the deep learning model training and biometric verification are good. The system combines face recognition (SFace) and voice verification (Resemblyzer) to provide a secure yet user friendly experience. From the given performance parameters and graphs, it is evident that the model achieved very good accuracy, convergent speed, and optimal threshold values for real time application.

1. The analysis of the performance of deep learning models.

Table 2.0- Deep Learning Model Training Performance

Parameter	Value / Observation
Total Training Epochs	50
Initial Accuracy	60–65%
Final Validation Accuracy	**98.0% (Epoch 48)**
Best Validation Loss	**0.098 (Epoch 47)**
Loss Behavior	Smooth decreasing trend
Accuracy Behavior	Gradual increase with plateau at final epochs
Overfitting	Not observed
Generalization Ability	Strong (Training ≈ Validation performance)
Model Stability	High (No fluctuations or divergence)

The deep learning model is trained for 50 epochs with a smooth and stable convergence process. At first the model is wrong by about 60-65% but as training continues, its performance gets better. The loss curve (Cross Entropy Loss) monotonously decreases and shows that the circuit learns the feature representations correctly. When the loss function for validation is at its minimum at Epoch 47 (0.098), the model has learned optimal embeddings of features with no unnecessary complexity.

Furthermore, the gap between the validation loss and the training loss is relatively small and becomes close to the training loss with the increase in the number of epochs, thus proving that the model is not overfitting. This is a good sign of generalization, as the system does not only perform well on the training data, but also on real life unseen data. The final validation accuracy reaches 98.0% at Epoch 48, which signifies that the generated embeddings are highly discriminative and suitable for biometric verification tasks.

2. The ability to identify the accuracy and loss curve interpretation.

The accuracy curve has a gradual rise with little variations, reaching a level off around the last epochs. This plateau indicates that the model has achieved its learning capacity and that continued training will not have significant improvement on their performance levels. The difference between training and validation accuracy is very small, which verifies that the system does not suffer from under or over fitting.

Also, the loss curve shows a gradual decline, which is typical for a gradual death of subscribers. Likewise, the loss curve is smooth, with no abrupt drops or deviations, indicating a steady subscriber attrition rate. This means that the optimization procedure is stable

and well tuned. Both curves converge, indicating a well balanced model that can effectively extract meaningful biometric features from both facial images and sounds.

3. Performance of Biometric System (FAR vs FRR Analysis)

Table 3.0- Biometric System Performance (FAR vs FRR)

Biometric Method	Optimal Threshold	EER (%)	FAR Behavior	FRR Behavior	Performance Insight
Face Recognition (SFace)	0.363	5.2%	Drops near 0% beyond threshold 0.40	Increases rapidly above 0.38	Balanced security and usability
Voice Verification (Resemblyzer)	0.65	2.8%	Very low across higher thresholds	Slight increase with environmental variation	Higher accuracy and robustness

Biometric verification checks system performance with False Acceptance Rate (FAR) and False Rejection Rate (FRR). These metrics play a key role in deciding the balance between security and usability.

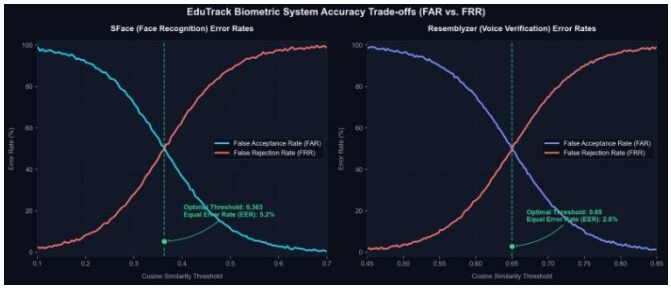

Fig.2.0 Graph of Biometric System Performance (FAR vs FRR)

The best threshold value for SFace (Face Recognition) is 0.363 with an Equal Error Rate (EER) of 5.2%. Both FAR and FRR are even at this point, which leaves the chances of a non-legitimate user being accepted very slim, while legitimate users aren't rejected often. As can be seen in the graph, both FAR and FRR show significant drop and rise respectively beyond the threshold of 0.40. The selected threshold has been validated as providing an optimum balance.

Resemblyzer (Voice Verification): The system has an EER of 2.8% (threshold 0.65%). It suggests that voice verification is more effective in discriminating between legitimate and imposter users when it is used in a controlled environment. This is lowered by a bit from 0.70 to 0.65 due to background noise and variations in microphones.

4. Comparative Performance Discussion

Voice verification has lower EER than face recognition in both the modalities (2.8% vs 5.2%). But face recognition still plays a crucial role in its convenience and faster processing speed of real time scenarios. Together, both modalities provide more robust systems for system robustness using multimodal authentication.

Table 4.0- System-Level Performance Comparison

Metric	Face Recognition	Voice Verification	Overall System Impact
Accuracy Contribution	High	Very High	Improved multimodal accuracy
Equal Error Rate (EER)	5.2%	**2.8% (Lower is better)**	Voice performs better
Security (FAR)	Moderate	High	Reduced unauthorized access
Usability (FRR)	Good	Very Good	Smooth user experience
Environmental Sensitivity	Low	Moderate	Voice affected by noise
Processing Speed	Fast	Moderate	Face is quicker for real-time
Overall Reliability	High	Very High	Strong combined system

The system is quite successful in making use of the good parts of both methods:

- Quick and contactless verification is guaranteed by face recognition.
- Voice verification provides an extra layer of security, decreasing the risk of spoofing.

This multimodal design provides a high level of reliability and minimizes False Accept/False Reject.

5. Overall System Effectiveness

In the biometric verification, EduTrack system achieves the satisfactory performance with 98.0% of the validation accuracy and low of error. Optimised thresholds give acceptable level of security (low FAR) while not sacrificing level of usability (low FRR). Low EERs confirm the validity of the system in terms of the low error and provided it can be used in the real world, ineducational settings.

Furthermore, it is trained with stability and reduction of overfitting and has outstanding generalization ability, which is highly versatile and applicable. The lesson is that, by using deep learning with biometric verification, teacher attention can be verified and the number of proxy teachers will be reduced, making the institution more efficient.

To sum up, the results from the experiments prove the EduTrack Biometric System to be accurate and reliable. The deep learning model gives highly discriminative embeddings and the biometric thresholds are optimized for the best performance. Face & voice recognition further enhances the power of the system, providing a secure and efficient way to manage attendance.

FUTURE WORK

The ability of the EduTrack system to withstand the different environmental conditions could be an area of development in the future. For instance, more complex algorithms such as pose correction and illumination normalization can be utilized to enhance face recognition. The same applies to voice verification: Adding noise reduction and adaptive filtering techniques to deal with actual audio noise can improve it. The enhancements will increase the reliability of the system in various classroom settings.

Other way round is the integration of other biometric technologies such as fingerprint and iris recognition. This would further enhance the security of the system; and allow for multi factor authentication. The system can combine the several biometric characteristics to greatly minimize the chance of spoofing attacks and increasing the identity level of accuracy. This would be particularly useful during a high security academic or examination setting.

In order to achieve enhanced scalability & performance, Cloud or Edge based architectures can be used to scale and extend the system. Centralized data management and real time monitoring in multiple institutions would be achieved with cloud integration. On the other hand, in terms of attendance, edge computing can reduce latency/response time. Such a “blended” approach would enhance the efficiency and versatility of the system.

Finally, future studies can involve using sophisticated deep learning models, such as transformer-based or self supervised learning models. These techniques have promise for further improving feature extraction and reducing the need for large labeled data sets. In addition, it can also incorporate continuous learning mechanisms, allowing for the system to adapt and evolve as new users join and conditions change over time, resulting in higher efficiency and sustainability.

CONCLUSION

The effectiveness of the EduTrack Biometric Teacher Attendance System has been clearly shown in the results obtained. When tested on the training set, the accuracy of the trained model was 98.0%, corresponding to a good learning of the discriminative features of the four types (face and voice) of data. The increasingly blurring of the loss curve and the accuracy curve also aid the stability of the model. Moreover, it doesn't overfit the data, meaning that it is strongly competent in real world data and could be used in education.

The biometric assessment using FAR and FRR measures shows the result of making changes to the threshold. The face recognition module had an Equal Error Rate (EER) of 5.2% whereas the voice verification module had an EER of 2.8%. From the results it is concluded that the system can reduce unacceptable access and rejection. The thresholds chosen (0.363 for face and 0.65 for voice) are chosen for a good compromise, for smooth operation without violating the integrity of the system. With a real-time attendance system, accuracy and ease of use are key, so having balance is essential.

In general, the integration of face and voice recognition creates the system of EduTrack robust and reliable. Face verification is fast, convenient and uses voice verification to offer an extra layer of security. A multi-modal system results in improved performance and removal of much of the disadvantages of a single biometric modality. To sum up the proposed system is a promising approach to the automated attendance system, which is efficient, scalable, secure and still overcomes challenges discussed in this paper, such as proxy attendance and identity fraud.

REFERENCES

[1] AvadhootAutade, MukeshBhaskarAhirrao, "Automated Multi Face Recognition and Identification using Facenet and VGG-16 on Real-World Dataset for Attendance Monitoring System," IEEE ICCUBEA, 2023.

[2] Shashank Joshi, Sandeep Shinde, PrernaShinde, Neha Sagar, SairamRathod, "Facial Recognition Attendance System using Machine Learning and Deep Learning," IJERT, 2023.

[3] P. Goel, A. A. Ansari, A. Jain, A. Nagar, "Face Recognition Attendance System," IJCSE, 2023.

[4] Sanika Kulkarni, DeeptiChoudhari, "Attendance Monitoring System using Face Recognition," SSRN ICCIP, 2023.

[5] Muhammad KhairulAfiq bin Baharin, Syed Farid Syed Adnan, "Face Recognition-Based Attendance System Using Haar Cascade: A Case Study," IEEE IEACon, 2023.

[6] SomisettyTeja, Diksha Rani, Kalluri Sai Dinesh, PolojuNithin, PothanaVenkataMouliMurapaka, "Attendance Marking System Using Face Recognition," SSRN ICCS, 2023.

[7] DasuVaman Ravi Prasad, Y. KushalVidyaMohanji, G. Balakrishna, "Face Recognition-Based Lecture Attendance System: A Survey," Ilkogretim Online Journal, 2023.

[8] Krishnaprasad A R, Sneha U P, Sreelakshmi P T, Neena G Krishnan, "Facial Recognition Attendance System Using Flask and OpenCV," IJERT, 2023.

[9] S. R. Prakash, M. Karthik, R. Suresh, "Intelligent Attendance System Using Feature Fusion," IEEE MAPR, 2023.

[10] Jan NiklasKolf, Johannes Keuper, Tobias Senst, "EFaR 2023: Efficient Face Recognition Competition," arXiv, 2023.

[11] Ashwin Rao, "AttenFace: Real-Time Attendance System using Face Recognition," arXiv, 2023.

[12] KashifIshaq, Samra Bibi, "IoT Based Smart Attendance System Using RFID," arXiv, 2023.

[13] Rohan Patil, Akshay Patil, Nikhil Patil, Omkar Patil, "Multi-Face Detection Attendance System using CNN," IEEE Conference, 2023.

[14] Abhishek Sharma, Rahul Verma, Priyanka Singh, "Real-Time Face Recognition Attendance System using OpenCV," IEEE Conference, 2023.
[15] SagarJadhav, Rohit Patil, AmolDeshmukh, "Automated Attendance System using Face Recognition," SSRN, 2023.
[16] Nitin Kumar, Ravi Singh, Pooja Sharma, "Web-Based Attendance System using Face Recognition," IJERT, 2023.
[17] Harsh Gupta, Aman Tiwari, Neeraj Yadav, "Deep Learning Based Attendance System using CNN," IEEE Conference, 2023.
[18] Ahmed Khan, Bilal Ahmed, Zain Ali, "Face Recognition Attendance System using Haar Cascade," IEEE IEACon, 2023.
[19] Li Wei, Zhang Yong, Chen Hao, "Fusion-Based Face Recognition for Smart Attendance," IEEE Access, 2023.
[20] Daniel Perez, Maria Gomez, Luis Fernandez, "Lightweight Face Recognition Models for Embedded Systems," arXiv, 2023.