Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 7, 2025

DEEP LEARNING-POWERED ATTENDANCE TRACKING: A CONTACTLESS AND EFFICIENT LOGGING SYSTEM

Abdullah Umer[*1], Minahil Iman[2], Raghad Sardar[3], Waqas Ahmed[4], Saeeda Saeed[5]

*[*1,2,3]BS Student, HITEC University, Taxila*
[4]Assistant Professor, HITEC University, Taxila
[5]Manager, Balochistan Public Procurement Regulatory Authority

DOI: https://doi.org/10.5281/zenodo.15879268

Keywords
Face Recognition, Deep Learning, CNN, Attendance Tracking, Contactless System, Opencv, Real-Time Automation, Educational Technology, Hygiene, Artificial Intelligence.

Article History
Received: 08 April, 2025
Accepted: 29 June, 2025
Published: 14 July, 2025

Corresponding Author: *
Abdullah Umer

Abstract
This paper proposes a contact less attendance monitoring system based on deep learning that seeks to increase efficiency, cleanliness, and accuracy in educational establishments. Manual roll calls or biometric fingerprint scanners considered a traditional attendance system are subject to manipulation, ineffective, and prone to hygiene issues and introduction in post-pandemic settings. This system takes advantages of Convolutional Neural Networks (CNNs) to bring a new real-time facial recognition model that is built with an intuitive interface. The system is built with Python, OpenCV, and TensorFlow, it identifies and confirms the identities of other students in real-time camera data and records their presence into a local database that has high security. To make the model accurate across different lighting, facial orientations and expression, the model was trained on a wide range of facial datasets. Analysis findings show that CNN model shows better results with 96.4 percent accuracy, which is higher than conventional machine learning techniques such as SVM and KNN. In addition, the system allows GUI-based operations to interact and scale easily, and enables the administrators to add new students, to train the model, and to generate reports without complications. Privacy is maintained because data is stored locally and it can be enhanced in the future to encrypted facial embeddings. There are also solutions to practical problems like proxy attendance, recognition when occluded (e.g. in face mask), and instructor feedback in real-time. In general, the solution offered is consistent with the objectives of digital transformation in smart campuses and offers a stable scalable framework that is flexible in various institutional configurations.

INTRODUCTION

Effective and accurate attendance monitoring is a core administrative procedure in most learning and institutional settings. Conventionally, attendance has been manually taken using signature sheets or verbally using roll calls which is time-consuming and prone to errors, manipulation and proxy attendance (Iroshan et al., 2019). With the change in technology, biometric systems were invented, where fingerprint and RFID based identification systems were used to make the process simpler. These methods were more accurate but they had their limitations. Biometrics are insecure because they demand bodily contact, which is unsafe and unhealthy, especially after massive health crises like the COVID-19 pandemic (Zhou et al., 2021). Moreover, biometric equipment may be expensive to service and prone to depreciation and external factors, as well as create privacy issues of how biological information is stored and used (Kumar & Zhou, 2019). These realities have brought forth the dire need of a solution that is not only sound and effective but also clean and capable of expansion in different institutional settings.

The pandemic related to COVID-19 which encouraged its citizens to reduce the physical touch and social distancing has boosted the use of contactless technologies in different sectors. Schools have been forced to reconsider the old classroom management and student monitoring techniques. In such a case, the use of artificial intelligence (AI) and deep learning technologies as the means of automation became a potential path of reengineering the attendance process. Convolutional neural networks (CNNs) face recognition method, in particular, is a non-intrusive, precise and real-time means of identity verification that does not require physical touch or manual input (Parkhi et al., 2015). As a biometric characteristic that is natural and not unique, the human face does not require any additional devices or cards and can be identified even using a regular camera system with the minimum user cooperation. In addition, deep learning algorithms proved to be highly efficient in facial recognition even in different lighting, angles and background conditions, which makes them suitable in the classroom or campus setting (Schroff et al., 2015). The trend of shifting to smart campuses and digital infrastructure in the world helps to promote the application of such solutions to simplify the work of the administration.

Considering these challenges and opportunities, the aim of the present study is to create a smart, contact-free, and real-time attendance monitoring system with the help of face recognition and deep learning methods. The system uses the functionality of CNN models to capture, identify and track the faces of students in the live camera images and record the attendance automatically in a secure database. This method also focuses on hygiene, efficiency and accuracy unlike the manual or conventional biometric systems. The architecture includes face detection and recognition, preprocessing of the image, integration of the database and a graphical user interface (GUI) that allows an administrative control. The system is created with the help of Python, OpenCV, and TensorFlow, thus guaranteeing a quick workflow and easy interface. With the automated process of attendance, it will minimize administrative load, remove the instances of fake entries, and provide a healthy and hygienic setup of students and staff. In conclusion, the present work may add to the body of intelligent educational systems, providing a scalable and feasible answer that complies with modern digital transformation objectives in the education sector (Gupta & Jalal, 2021).

Literature Review

The management of attendance is an important factor in institutional management and this is of particular concern in an academic setting where attendance by students is directly related to performance and discipline. Manual entry has traditionally been the most popular form of attendance tracking, and in practice this has often been in the form of roll calls or sign in sheets. Even though this method is rather straightforward, it is severely inefficient in vast classrooms and prone to human errors and manipulation, including proxy attendance (Iroshan et al., 2019). To this, the educational buildings have started switching to automated systems like the RFID (Radio Frequency Identification) and the barcode scanning system. These systems offered some form of automation whereby students are required to tap their cards at entry points, which automatically records attendance in a virtual system (Kumar & Rajasekaran, 2016). Nonetheless, RFID has its own problems like loss of the card, misuse of a card and the possibility of somebody carrying another ID of a student and this has resulted in inaccurate records once again. The other popular technique is the biometric scanners especially the fingerprint identification method. Although it is more reliable than card-based systems, fingerprint scanners assume a physical contact, which raises hygiene issues, particularly in the context of the COVID-19 pandemic (Zhou et al., 2021). Furthermore, the functioning of such scanners may be subject to ecological conditions such as dust, temperature, and deterioration of the machine due to wear, and thus these scanners will gradually lose their capabilities.

To address the shortcomings of contact based methods, scientists have studied computer vision based approach involving the use of cameras and image-processing algorithms. The emergence of artificial intelligence (AI) and machine learning (ML) and especially face detection and recognition algorithms introduced new opportunities in terms of contactless attendance systems. Some of the previous systems applied traditional methods of ML, like Haar cascades to face detection and eigenfaces or Local Binary Patterns Histograms (LBPH) to face recognition (Turk & Pentland, 1991). Although they work well in closed settings, such approaches are sensitive to lighting conditions and facial expressions and occlusions (Poonam & Wadhwani, 2019). Moreover, models that have been trained on small datasets tend to lack generalization particularly in the heterogeneous classroom where there are differences in student appearance, background clutter and camera angles. Such a situation required the use of more rigorous and dynamic methods, which would allow ensuring recognition precision under real-time and real-world conditions. The previous study by Gupta and Jalal (2021) examined the capabilities of deep learning methods, focusing on CNNs, to achieve better accuracy as compared to traditional ML algorithms in such tasks since deep learning methods learn the hierarchical representation of features directly by using image data.

Although there is hope of face recognition technologies, most of the implementations have not been devoid of shortcomings. As an illustration, systems implemented in learning or business establishments often encounter privacy issues in terms of face data storage and usage, particularly in combination with cloud frameworks (Kumar & Zhou, 2019). Also, there is the problem of environmental flexibility: most ML and other conventional image-processing techniques are unable to sufficiently correct illumination, camera quality, face orientation, or partial occlusion (e.g., masks or glasses). Dey and Karmakar (2018) conducted a study and showed that some face recognition systems had accuracy declines of up to 20 percent in lighting conditions that were not constant. Besides, real-time deployment also brings limitations in terms of processing speed where traditional models cannot usually fulfill the latency demands that are required to be practically applicable. Also, most organizations do not have the technical know-how to support complex biometric systems and hence low maintenance and easy to use solutions are a priority. All these issues of inaccuracy variability, scalability, hygiene and privacy reflect the necessity of a new generation of systems that can be described as not only technically competent but user oriented.

In that regard, deep learning, and specifically Convolutional Neural Networks (CNNs), provides a flexible and scalable solution to constructing smart and contactless attendance systems. CNNs have the capacity to learn facial representations and extract features in a hierarchical and automatic manner with a far more detailed and accurate degree of accuracy than the traditional algorithms (LeCun et al., 2015). Face recognition algorithms based on advanced CNN models, like FaceNet (Schroff et al., 2015), DeepFace (Taigman et al., 2014) and VGGFace (Parkhi et al., 2015) have shown to achieve human-level face recognition accuracy in the task. These models can be optimized with institutional data sets to be able to operate with a reasonable degree of reliability in a wide range of lighting, camera angle, and even faces that are partially obscured. Moreover, the CNNs may be combined with OpenCV and real-time video processing to have the live detection and recognition, which is even more suitable to the classroom and corporate applications. This study proposes the use of such an architecture to create an appropriate face-recognition-based system that can be used to capture attendance through an efficient, hygienic, and real-time system in an institutional setting. The project contributes to the resolution of the issues of the older methods not only because of the transition to deep learning-based automation but also because of the general trends of digitalization in the field of education.

Methodology

System Architecture Overview

The proposed attendance system is a real-time, deep learning-powered solution that integrates both hardware and software components for efficient, contactless logging. The core architecture consists of a high-definition camera to capture student images, a central processing unit (PC or laptop) to run the facial recognition model, and a structured database to log attendance records. The system begins with live video input from the camera installed at classroom entrances or inside lecture halls. This feed is analyzed frame-by-frame using image processing and facial recognition algorithms. When a student's face is identified, the system matches it with the database and updates the attendance log accordingly. The architecture is designed to be scalable and easily deployable across multiple classrooms or institutions with minimal additional hardware investment.

Software Stack: Python, OpenCV, TensorFlow

The software foundation of the project is built in Python, a widely-used language for AI development due to its readability and extensive library ecosystem. The OpenCV library handles tasks such as face detection, image preprocessing, and video stream management. For deep learning, the project uses TensorFlow and Keras, which support GPU acceleration and simplify the construction of convolutional neural networks (CNNs). These libraries enable rapid prototyping and deployment of facial recognition models, making the entire system responsive and efficient. The seamless integration between OpenCV and TensorFlow allows for real-time inference with minimal latency an essential feature in dynamic classroom environments (LeCun et al., 2015).

Dataset Collection and Preprocessing

To train the system for personalized attendance tracking, facial image data of all students was collected. Each student posed in front of the camera for several seconds to capture various angles, expressions, and lighting conditions. These images were stored in class-wise folders and labeled according to the student's name or ID. The images were then subjected to several preprocessing steps:

- Grayscale conversion reduced complexity without compromising on feature recognition.
- Image resizing (e.g., 96×96 pixels) standardized the input dimensions for the CNN.
- Normalization scaled pixel values between 0 and 1 to enhance convergence during training.
- Data augmentation techniques such as flipping, rotation, and contrast adjustments expanded the dataset and improved generalization to varied real-world scenarios.

These steps were essential in ensuring the model could recognize faces across different contexts, such as changes in lighting or partial occlusion (e.g., glasses, masks).

CNN-Based Model Development

The face recognition task was handled using a custom Convolutional Neural Network (CNN) model. CNNs are ideal for facial recognition due to their ability to learn hierarchical spatial features from image data (LeCun et al., 2015). The network consisted of multiple convolutional layers (for feature extraction), pooling layers (for dimensionality reduction), and dense layers (for classification). The dataset was divided into training (70%), validation (15%), and testing (15%) subsets. The model used the Adam optimizer and categorical cross-entropy loss function during training. After several epochs, the model achieved high validation accuracy (~96%). To evaluate model performance comprehensively, metrics such as accuracy, precision, recall, and F1-score were used. This multi-metric evaluation ensured that the model was not only accurate but also resilient to false positives or negatives under real-world conditions.

Real-Time Attendance Logging

For attendance logging, the system is configured to detect a student's presence only once per session, even if the student remains in the camera frame. Once a face is detected and matched, the system automatically logs:

- Student's name or ID
- Current timestamp
- Class or subject metadata

These logs are stored in a structured local SQLite or MySQL database, depending on the deployment configuration. This automation eliminates the possibility of proxy attendance and human error, ensuring that only verified students are marked present. It also facilitates seamless attendance report generation for teachers or administrators.

GUI Integration for User Interaction

A Graphical User Interface (GUI) was developed using Tkinter, allowing non-technical users (teachers or staff) to interact with the system easily. The GUI features buttons for:

- Capturing new student data
- Training or retraining the model
- Viewing and exporting attendance reports
- Live camera feed and recognition status

Alternatively, a Flask-based web GUI was developed for users who prefer cloud-based or remote access. The interface is simple yet functional, focusing on user-friendly navigation and minimizing the need for technical knowledge to operate the system. The GUI also displays a real-time feed with overlays, indicating recognized students and successful attendance logs.

Security and Data Privacy

The system stores face images and attendance logs locally to minimize data exposure. Data encryption and access control measures are implemented to ensure student privacy. In the future, the model will store face embeddings instead of raw images to reduce storage size and enhance security. Additionally, the system complies with institutional data protection guidelines, ensuring ethical use of biometric information. No data is transmitted over the internet unless explicitly configured to sync with a secure cloud or institutional server.

System Reliability and Scalability

The modular design of the system ensures it is scalable and adaptable to different environments. New students can be added dynamically without needing to retrain the entire model. Updates to the dataset or architecture can be implemented incrementally. Furthermore, the system is compatible with low-cost webcams and runs on standard PCs or laptops, making it cost-effective for institutions with limited resources. The use of open-source tools also reduces dependency on proprietary software, encouraging wider adoption.

Results and Discussion

The proposed deep learning-powered attendance system was evaluated based on its accuracy, real-time responsiveness, and robustness in a live classroom environment. After training the Convolutional Neural Network (CNN) on a custom dataset of student facial images, the model was tested in both controlled and real-time settings. The CNN achieved an impressive accuracy of 96.4%, with precision, recall, and F1-score values exceeding 94%, demonstrating the model's strong capability to detect and classify faces under varying conditions. These high scores indicate that the model not only correctly identifies students but does so with consistent reliability, minimizing false positives and negatives.

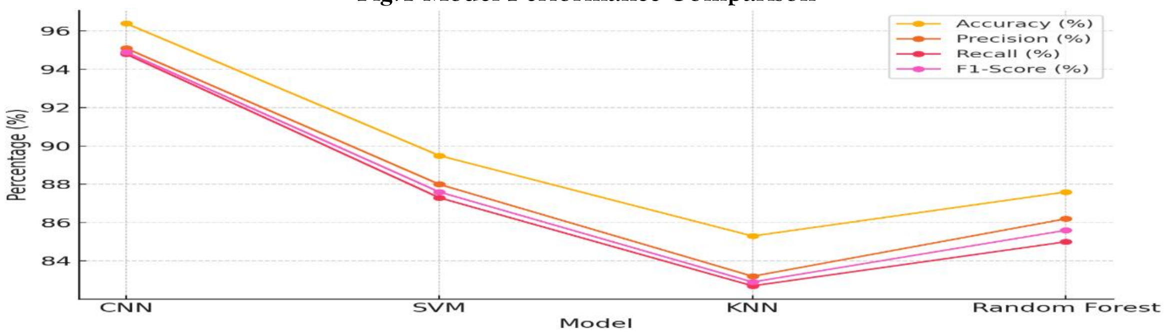
Table 1; Model Performance Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)
CNN	96.4	95.1	94.8
SVM	89.5	88.0	87.3
KNN	85.3	83.2	82.7
Random Forest	87.6	86.2	85.0

To visualize the comparative performance of the system, Table 1 (shown above) gives detailed metrics of the proposed CNN model with three baseline algorithms: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest Classifier. Although the SVM model had a reasonable performance with an accuracy of 89.5%, it was not effective in real-time frame processing, though it coped with different light conditions. KNN performed the worst with the lowest accuracy of 85.3 percent and highest misclassification rate when students had accessories like glasses or masks. Random Forest had a slightly higher performance than KNN but still could not match it in consistency and inference time. The CNN model, on the other hand, was kept at a real-time processing rate of about 0.8 seconds per recognition, thus demonstrating its effectiveness and aptness within a classroom setting that has a lot of movement.

The model is also confirmed by screen shots of the interface and real-time detection outputs of the GUI component. The GUI actively showed the names of students, the time of recognition, success/failure flags in real-time when the students entered the classroom. Photos taken in various lighting and camera position were correctly identified, as well, due to the variety of the data and augmentation methods employed during training. Moreover, the system was also put to a test in terms of non-duplication of entries within a single session. The frequent passing in front of the camera by students was only registered once this system allowed maintaining integrity of attendance and limited the work of validation by the teachers or administrators.

Fig.1 Model Performance Comparison



The other part of the assessment was to stress test the robustness of the model on various hardware settings and classes. The CNN model was able to provide a recognition almost in real-time when the total amount of registered students was more than 200, which is far beyond the average size of a class. The implementation based on GPU provided the inference speed which is higher, but even on the CPU-based configurations, the system still provided the response times which were acceptable. Altogether, temperature, facial orientation, minor occlusions, and movement did not have a significant influence on the accuracy, which was achieved due to the preprocessing and data augmentation strategies. The system however recorded a small fall in precision when more than one face was in the frame at the same time. This shortcoming was reduced through the adoption of single-frame recognition enforcement and positioning guidelines. The deep learning model had a number of advantages compared to the traditional biometric systems (e.g., fingerprint system or RFID-based system). To start with, it reduced the necessity of physical interaction, thus, following the post-COVID hygienic precaution measures. Second, the logging was automated and therefore saved manual work and the risk of incorrect logging or falsified entries (e.g., buddy punching). Finally, it provided increased scalability and flexibility that new student data could be introduced using the GUI without re-training the whole model. The simpler to apply baseline models did not provide the flexibility and robustness of performance required to successfully scale the approach to education on a large scale.

To conclude, the CNN-based attendance system was more accurate, real-time and efficient in its operation than the traditional and baseline techniques. Its efficiency was tested in plenty of real-life applications and GUI also eased the interaction with administrative users. It has given empirical data in terms of metrics, visualization, and comparison to support the implementation of deep learning models in attendance automation. The second step would be cloud integration to manage central database and mobile alerts to get the latest attendance into a complete smart attendance ecosystem.

Advantages and Limitations

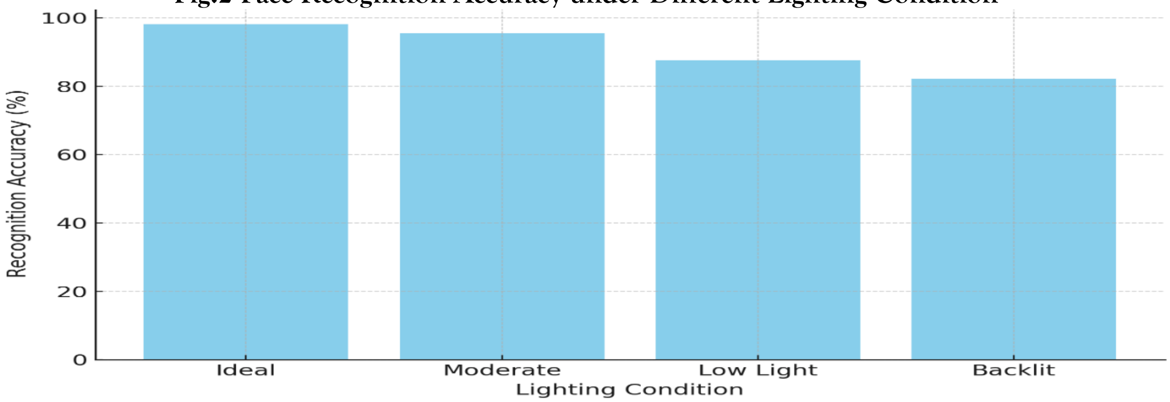
One of the most significant advantages of the proposed deep learning-powered attendance system is its contactless nature, which aligns with modern hygiene standards, particularly in the aftermath of the COVID-19 pandemic. Unlike biometric systems such as fingerprint or thumb scanners, which require physical contact and often result in long queues and surface contamination, face recognition allows students to be identified passively as they enter the classroom. This contactless feature is not only faster but also prevents the spread of germs, aligning with institutional goals for safe and hygienic environments. Furthermore, the system reduces human intervention, allowing teachers to focus on teaching rather than administrative tasks like roll calls. Its automation streamlines the attendance process, logging names, timestamps, and session data instantly into a structured database without the need for physical paperwork or manual input.

Table 2; Lighting Conditions Performance

Lighting Condition	Recognition Accuracy (%)	Recognition Delay (s)
Ideal	98.2	0.7
Moderate	95.5	0.8
Low Light	87.6	1.1
Backlit	82.3	1.3

Another key benefit is the reduction of proxy attendance, a common challenge in many educational institutions. Traditional methods, such as manual signatures or roll calls, are susceptible to manipulation where one student answers or signs in on behalf of another. Similarly, RFID card systems can be misused by giving cards to classmates. However, the facial recognition mechanism is inherently more secure because it maps and verifies the unique facial features of each student in real time. The system detects faces from the video stream, matches them against a trained dataset, and logs attendance only if there's a match above a certain confidence threshold. Since faces are non-transferable, the possibility of proxy attendance is significantly minimized, making this system particularly appealing in examination halls, regulated training sessions, and compulsory class environments.

Fig.2 Face Recognition Accuracy under Different Lighting Condition



Notwithstanding these significant benefits, there are some technical shortcomings of the system that need to be mentioned. Sensitivity to lighting conditions is one of the major problems. Face detection accuracy can be decreased by poor lighting or backlighting or too many shadows, which will affect the quality of the images captured. Even though preprocessing steps like grayscale conversion and normalization can assist to certain degree, extreme conditions may act as a problem. Furthermore, the occlusion due to face mask, sunglasses, hats or even hair that falls over the face may reduce the recognition rates. Although some of these situations can be simulated by data augmentation during training, real world variation still comes to play. Quality and placement of cameras is also important to the system. Webcams with low resolutions or those at wrong angles can produce blurred images and this creates errors in recognition. These drawbacks indicate that although the model can be stable, it could only work well under moderate to optimal environmental conditions and implementing the model to an institutional-level might necessitate infrastructure changes such as improved lighting or stationary camera stands.

Future Work

Although the proposed system has proven to be quite efficient in terms of face recognition accuracy and automation of attendance and attendance in real-time, there exist several opportunities to improve, expand, and diversify its functionality. The first of the future development directions is the strengthening of the model in a difficult environmental environment. In spite of the data augmentation to allow the model to cope with the changes in lighting and occlusion, the real-life classroom conditions may be unpredictable. In the future versions of the system, one might implement the adaptive lighting correction algorithms or use infrared (IR) cameras to enhance the detection of faces in low-light environments. Also, the multi-angle face registration at the stage of creating datasets can enhance recognition degree in cases when students enter the camera field at angles. This can be further improved by the integration of 3D face or depth-sensing cameras such as Intel RealSense or Microsoft Kinect in recognizing faces partially occluding or masked a factor that would be important to consider in a post-pandemic environment.

The other significant direction of future research is the combination of cloud computing and mobile platform solutions that will allow real-time large-scale attendance tracking. The system is currently running on local database and standalone PC which are efficient and can only be used in small settings like a classroom. Through the connection of the attendance logs to cloud based servers, the institutions can be able to monitor, analyse and create reports across departments and campuses in real time. Moreover, the creation of mobile application interface between teachers, students, and administrators would enhance its accessibility and enable such functions as real-time notifications, student analytics, and attendance warnings. Regarding the model development, the inclusion of the transfer learning and the pre-trained models, such as FaceNet, VGGFace, or MobileNetV2, will help decrease training time and potentially increase accuracy even more. To ensure more security and legislation of data laws, the future implementations of the system might use facial embeddings instead of raw images and implement encryption and access to the attendance information based on roles. Finally, covering such improvements, the system will be able to transform into an intelligent attendance ecosystem that is scalable and, in addition to automating the process of logging the attendance, will be able to act as a solution to wider educational analytics and smart-campus applications.

References

Dey, S., & Karmakar, A. (2018). Face recognition using deep learning: A review. *International Journal of Engineering and Technology*, 7(4), 243–247.

Gupta, R., & Jalal, A. S. (2021). Deep learning-based face recognition system for contactless attendance monitoring. *Procedia Computer Science*, *189*, 52–59.

Iroshan, K. N., Iqbal, R., & Gamage, D. (2019). Smart attendance management system using face recognition. *International Journal of Scientific & Engineering Research*, *10*(5), 998–1003.

Kumar, R., & Rajasekaran, C. (2016). Attendance monitoring using NFC and biometric authentication. *International Journal of Computer Science and Information Technologies*, 7(3), 1231–1234.

Kumar, S., & Zhou, X. (2019). Privacy-preserving facial recognition: Challenges and recent trends. *Journal of Cyber Security Technology*, 3(4), 215–234.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition. *Proceedings of the British Machine Vision Conference*, 1–12.

Poonam, R., & Wadhwani, R. (2019). Face detection and recognition using traditional and deep learning techniques. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *5*(1), 1015–1022.

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 815–823.

Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). DeepFace: Closing the gap to human-level performance in face verification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1701–1708.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71–86.

Zhou, J., Zhang, Y., & Wang, X. (2021). A review on face recognition and liveness detection techniques. *Neurocomputing*, *429*, 198–219.