International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Smart Attendance and Behavioural Analysis

Praneeth[1], Naresh[2], Sai Kiran[3], Anvesh[4], Deepak[5]

Under the guidance of **Mr.K.Y.Vamsi Naidu**(Asst. Professor)

Department of Computer Science and Engineering(AI & ML),Raghu Institute of Technology

Abstract: This project introduces the Smart Attendance and Behavioural Analytic System, designed to enhance student engagement and learning outcomes in modern educational environments. Utilizing advanced voice recognition technology, the system automates attendance marking, allowing students to check in simply by stating their name or a designated phrase.

This automation significantly reduces the time and effort required for manual attendance procedures. The system incorporates Natural Language Processing (NLP) and Machine Learning (ML) techniques to analyze students' vocal expressions, tone, and speech patterns, providing real-time insights into their emotional and engagement states during classes, such as attentiveness, confusion, and interest. Equipped with a user-friendly dashboard, educators can effortlessly track attendance, visualize engagement metrics, and identify behavioural patterns over time.

Equipped with a user-friendly dashboard, educators can effortlessly track attendance, visualize engagement metrics, and identify behavioral patterns over time. The system also issues alerts for drops in student engagement or persistent absenteeism, facilitating timely interventions. By integrating with Learning Management Systems (LMS), it streamlines attendance reporting and behavioural analysis, allowing institutions to monitor and support students effectively. The system also issues alerts for drops in student engagement or persistent absenteeism, facilitating timely interventions. By integrating with Learning Management Systems (LMS), it streamlines attendance reporting and behavioral analysis, allowing institutions to monitor and support students effectively.

I. Introduction

The modern educational landscape is rapidly evolving, and with advancements in technology, there is a growing opportunity to enhance both student engagement and the efficiency of classroom management. The project outlined here is designed to transform traditional attendance tracking and behavioral analysis in classrooms by integrating cutting-edge technologies such as voice recognition, natural language processing (NLP), and machine learning. By automating attendance through voice recognition, the system eliminates the time-consuming and error-prone manual attendance process, providing educators with accurate and real-time data on student presence. Beyond simply tracking attendance, the system analyzes students' vocal expressions, tone, and speech patterns to provide valuable insights into their emotional states, engagement levels, and behavioral patterns during class. This allows educators to assess attentiveness, confusion, interest, and other key indicators of student engagement, enabling them to make more informed, personalized decisions in real time. Furthermore, the system is designed to integrate seamlessly with existing Learning Management Systems (LMS), offering a streamlined solution that consolidates attendance, behavioral data, and engagement metrics in one platform. The project also aims to foster proactive intervention by issuing alerts when students show signs of disengagement or absenteeism, ensuring that educators can provide timely support. Ultimately, the goal is to create an environment where educators can leverage technology to enhance their teaching methods, address student needs more effectively, and ultimately improve both learning outcomes and classroom dynamics. Through a combination of automated systems, data-driven insights, and user-friendly interfaces, this project promises to revolutionize how schools track attendance and monitor student behavior, leading to a more responsive, efficient, and personalized educational experience.

II. Primary Objective

This project is designed to transform traditional classroom environments by enhancing student engagement, streamlining attendance management, and providing valuable behavioural insights to educators. By fostering an atmosphere that encourages active participation and improved academic performance, the system aims to redefine how educators interact with their students. The tedious and time-consuming task of manual attendance is eliminated through a fully automated voice recognition system, enabling students to check in by stating their name or a designated phrase with a high degree of accuracy to minimize errors. Beyond attendance, the system leverages advanced natural language processing (NLP) and machine learning (ML) techniques to analyze vocal expressions, tone, and speech patterns. This allows for the assessment of emotions such as attentiveness, confusion, and interest, providing real-time insights into student engagement that educators can use to make timely and informed decisions.

A comprehensive and user-friendly educator dashboard is central to the system, offering features to track attendance, visualize engagement metrics, and identify long-term behavioural patterns and trends. Alerts are issued automatically for critical engagement issues, such as significant drops in student participation or persistent absenteeism, empowering educators to intervene proactively. By integrating seamlessly with existing Learning Management Systems (LMS), the solution synchronizes attendance and behavioural data, simplifying reporting and monitoring processes for institutions while ensuring compatibility with current workflows.

Furthermore, the system facilitates personalized educational support by providing insights into individual student behaviour, helping educators tailor their teaching strategies to meet the specific needs of each learner. Whether it's addressing signs of disengagement, confusion, or absenteeism, the platform ensures that every student receives the attention and assistance they need to succeed academically and personally. Ultimately, this project aims to empower educators with the tools and data necessary to create an enriched, responsive, and efficient learning environment that drives better outcomes for both students and institutions.

III. Inputs required

To successfully implement the project, several critical inputs are necessary to ensure that the system operates efficiently and accurately in a real-world educational environment. First and foremost, **voice data** is essential for training and testing the voice recognition system. This includes a diverse collection of student voice samples, which will allow the system to identify and authenticate students based on their vocal characteristics. These voice samples should cover a range of phrases, including the student's name or a designated phrase used for check-in. Additionally, to ensure the system can handle various accents, speech patterns, and potential background noise, it is crucial to incorporate diverse voice datasets. This diversity ensures the system can maintain high accuracy and adaptability when processing speech from students with different linguistic and environmental factors.

Second, **student information** is vital for proper identification and attendance tracking. The system needs detailed student data, such as names, unique student IDs, and enrollment information, which will be used to match voice samples with the corresponding student records. This allows for precise tracking of student attendance, ensuring that the correct individual is marked present.

Next, **classroom environment data** plays a significant role in optimizing the system's performance. Each classroom has its own set of acoustic properties—such as background noise levels, room size, and echo effects—that can influence the clarity and accuracy of voice recognition. Therefore, it is important to have **acoustic profiles** for each classroom, which will help the system adjust to different environments and reduce errors caused by external factors. In addition, **real-time audio feeds** from the classroom are needed, which will continuously capture the students' speech for both attendance marking and behavioral analysis. This real-time data will allow the system to assess not only who is present but also to evaluate the emotional states and engagement levels of students based on their speech patterns during class.

Integration with existing school technologies is also crucial for seamless operation. The project requires access to the **Learning Management System (LMS) API**, which will enable the system to synchronize attendance and behavioral data with the school's existing platforms. This integration simplifies the management and reporting processes, allowing educators to access real-time attendance and engagement data in one unified interface.

Finally, the system requires **behavioral data**, specifically datasets of speech patterns associated with different student behaviors, such as whether a student is sitting, standing, moving, or showing signs of disengagement (e.g., speaking in a monotonous tone or speaking less). This data will be used to train machine learning models that can assess not only whether a student is present but also whether they are engaged, attentive, or potentially struggling. The ability to analyze these behavioral cues will provide educators with real-time insights into their students' emotional states, such as confusion or boredom, allowing for more personalized interventions and support.

Collectively, these inputs—voice data, student information, classroom environment data, LMS API access, and behavioral data—are fundamental to creating a comprehensive system that integrates voice recognition and behavioral analysis to improve student engagement, streamline attendance management, and empower educators with actionable insights for personalized interventions. These elements work together to ensure the system is both accurate and effective in enhancing the educational experience for both students and teachers.

IV. Software and Tools

To successfully implement the project, several software and tools are required to handle different aspects of the system. For **voice recognition**, tools like Google Speech-to-Text, IBM Watson, or custom voice recognition systems are essential for accurately transcribing and identifying student voices. For **Natural Language Processing (NLP)**, frameworks such as SpaCy, NLTK, or Hugging Face Transformers are needed to analyze speech for emotional cues and tone, helping assess student engagement. In terms of **machine learning**, frameworks like TensorFlow, PyTorch, or scikit-learn will be used to develop and train models for behavioral analysis, enabling the system to assess students' emotional states and engagement levels. A **Database Management System** such as MySQL or MongoDB is necessary to store critical data, including attendance records, engagement metrics, and behavioral data. Finally, **visualization tools** like D3.js, Tableau, or Plotly are crucial for building the educator dashboard, which will allow teachers to track attendance, visualize student engagement trends, and gain insights into behavioral patterns through intuitive, interactive interfaces. Together, these tools will create a comprehensive system for enhancing student engagement and improving classroom management.

V. Outcomes

The primary outcomes of the project include an automated attendance system that enables efficient, accurate, and time-saving voice-based attendance tracking, effectively eliminating manual errors in attendance records. This system also provides real-time behavioral insights, offering instant analysis of student engagement levels during classes. It allows educators to identify emotional states such as attentiveness, confusion, or disinterest, giving them valuable information to improve their teaching strategies. Additionally, the system enhances educator effectiveness by providing actionable insights that allow teachers to tailor their teaching methods to individual student needs. It also facilitates the early detection of disengagement or absenteeism, enabling timely interventions to ensure that students stay on track. Secondary outcomes include improved student learning outcomes, where better engagement strategies lead to higher participation, understanding, and academic performance. The system also offers targeted support for students struggling either academically or emotionally, addressing their specific needs through data-driven insights. Moreover, data-driven decision making is promoted as institutions can leverage attendance and behavioral data to improve academic policies and gain insights into teaching effectiveness and classroom dynamics. Seamless LMS integration is another key benefit, where attendance and behavioral data are centralized within existing LMS platforms, streamlining reporting processes for administrators and educators. The project also features user-friendly dashboards, providing comprehensive views for educators to track attendance trends, engagement metrics, and behavioral patterns over time. Finally, the system includes proactive alerts that notify educators and administrators about persistent absenteeism or low engagement, providing tools to ensure timely interventions and better student support. These outcomes collectively aim to create a more efficient, personalized, and data-driven educational environment that benefits both students and educators.

VI. Background Study

Facial Recognition based Smart Attendance System: The Facial Recognition based Smart Attendance System presented in the ISO 9001:2008 Certified journal (2024) offers a versatile architecture that can seamlessly integrate into various applications, including attendance systems. It utilizes advanced models such as RetinaFace-10GF and ResNet50 to enhance accuracy and efficiency. RetinaFace-10GF is a powerful face detection model capable of detecting faces with high precision under different poses, scales, and occlusions, even in challenging conditions such as poor lighting.[1] ResNet50, a deep convolutional neural network with 50 layers, improves feature extraction and recognition accuracy, particularly in image classification tasks. By leveraging these cutting-edge technologies, the system aims to reduce administrative burdens in educational environments while ensuring reliable attendance marking. The advantages of this system include its non-invasive and user-friendly nature, high precision in face detection, real-time attendance monitoring, and automation that reduces manual effort and proxy attendance issues. However, the system also faces challenges, including difficulties in uncontrolled environments with varying lighting and occlusions, reliance on a robust hardware setup for video capturing, and potential privacy concerns related to the storage of facial data.

Local Binary Pattern (LBP) Optimization for Feature Extraction: The paper titled "Local Binary Pattern (LBP) Optimization for Feature Extraction", published in the Indonesian Journal of Electrical Engineering and Computer Science in 2024, focuses on enhancing the use of texture as a critical feature in image processing tasks. Texture analysis is fundamental for image understanding, and the Local Binary Pattern (LBP) operator, which describes local texture features of images, plays a pivotal role in this domain. The paper presents an innovative approach to LBP by dividing the operator into three matrices, two of which are fixed and independent of the input data. These fixed matrices are thoroughly analyzed, and a new algorithm is proposed to optimize them, aiming to enhance classification performance. The optimization technique leverages the singular value decomposition (SVD) algorithm, which allows for the creation of optimal LBPs that significantly improve the texture description of human face images. Experimental results demonstrate the effectiveness and superiority of these optimized LBPs in tasks like face detection and facial expression recognition.[4] One of the key advantages of this system is the integration of Support Vector Machines (SVM) and LBP, which ensures robust recognition capabilities. Furthermore, it is designed to handle real-time attendance tracking in classroom settings, with a modular design that can be scaled across institutions. The system also shows high accuracy in controlled environments, making it suitable for various applications. However, there are some drawbacks, such as inefficiency when dealing with dynamic classroom conditions where frequent movement occurs, potential performance degradation under poor lighting or occlusions, and the computational overhead introduced by the feature extraction method. Despite these challenges, the optimized LBP approach offers significant improvements in face detection and facial expression recognition, making it a valuable tool in image processing applications.[12]

Applying Multi-task Cascade Convolutional Neural Network (MTCNN) for Face Detection: The paper titled "Applying Multi-task Cascade Convolutional Neural Network (MTCNN) for Face Detection", published in INFORMS PubsOnLine in 2020, explores the application of the Multi-task Cascade Convolutional Neural Network (MTCNN) for the task of face detection. MTCNN is a powerful approach that works in three distinct stages: the first stage, called P-Net, predicts candidate facial regions, the second stage, R-Net, filters the bounding boxes, and the final stage, O-Net, proposes facial landmarks to refine the detection process. To implement this approach, the study used 10 images, nine of which were obtained from the publicly available Real and Fake Face Detection dataset on Kaggle, and the last image was captured by photographer Bob N. Renee. The approach aims to provide an efficient and effective method for real-time face detection with multiple applications, such as attendance tracking in classrooms. The system offers several advantages, including the ability to mark attendance in real-time without the need for manual intervention, scalable database management through the use of MongoDB, and the utilization of the Haar Cascade algorithm for lightweight and efficient face detection. The integration of multiple cameras further reduces blind spots in classrooms, ensuring better coverage. However, there are certain disadvantages, such as potential accuracy issues when relying on the Haar Cascade algorithm in low-resolution or complex scenarios, challenges in distinguishing between identical twins or individuals with similar facial features, and the system's dependency on a consistent internet connection for real-time updates.[1] Despite these limitations, the MTCNN-based approach provides a promising solution for face detection, with potential benefits in automated attendance marking and other applications requiring accurate facial recognition.

Multi-RPN Fusion-based Sparse PCA-CNN Approach to Object Detection and Recognition for Robot-aided Visual System: The paper titled "Multi-RPN Fusion-based Sparse PCA-CNN Approach to Object Detection and Recognition for Robot-aided Visual System", published in IEEE in 2017, addresses critical challenges in robot-aided visual systems, particularly in the area of object detection and recognition[3]. Object detection plays a pivotal role in enabling robots to interact with their environment, but current algorithms, such as Fast R-CNN or Faster R-CNN, struggle with high misdetection rates and low computation speed when the object's scale varies significantly. Moreover, these methods suffer from the inherent problem of overfitting, which limits their generalization ability. To overcome these challenges, the paper proposes a novel approach that integrates a multi-RPN (Region Proposal Network) fusion method to generate candidate windows, followed by a sparse PCA-CNN (Principal Component Analysis-Convolutional Neural Network) algorithm for more accurate object detection and recognition from images or video sequences. The authors conducted experiments to validate the proposed algorithm, and the results showed improved performance compared to existing methods. This approach is particularly valuable in contexts where object detection and recognition are crucial for autonomous systems, such as robots.[5]The advantages of this method include its potential for contactless attendance tracking, which enhances safety in post-COVID scenarios by eliminating the need for physical interaction. The system can be integrated with existing educational platforms, offering seamless usability and leveraging CNNs for high-accuracy facial recognition. Additionally, it is robust against manual roll-call errors and proxy attendance, making it highly reliable. However, there are certain drawbacks, including the high dependency on robust camera setups and optimal lighting conditions, which may limit its effectiveness in less controlled environments. Furthermore, the computationally expensive models could hinder real-time performance in larger classrooms, and the training of these models requires substantial labelled data to achieve high effectiveness, which can be a challenge for large-scale implementations. Despite these limitations, the proposed approach offers significant advancements in object detection and recognition for robotic and automated systems, particularly in educational settings.[6]

Automated Face Recognition System for Smart Attendance Application Using Convolutional Neural Networks: The paper titled "Automated Face Recognition System for Smart Attendance Application Using Convolutional Neural Networks", published in IEEE in 2024, proposes a touchless automated face recognition system designed for smart attendance applications, particularly useful in offices and educational institutions. This system is particularly relevant in the context of the ongoing global health challenges, such as the spread of COVID-19, as it minimizes physical contact, helping to reduce the transmission of viruses. The system uses a **Convolutional Neural Network (CNN)**, trained with a dedicated database of 1,890 faces, incorporating various illumination levels and rotational angles across 30 targeted classes. A performance analysis of the CNN was conducted using both 9-layer and 11-layer networks with different activation functions, including **Step, Sigmoid, Tanh, Softmax**, and **ReLu**. The results showed that the 11-layer CNN with the **ReLu activation function** provided the highest accuracy of 96.2% on the designed face database. The system utilizes the **Viola-Jones algorithm** to detect multiple faces in test images, ensuring robust face detection.[2] Additionally, a web application was developed to monitor attendance and generate detailed reports, providing an efficient solution for automated attendance tracking. The advantages of this system include eliminating manual errors through **deep learning-based attendance** automation, the use of multiple robust algorithms like Viola-Jones and CNN for face detection, and its high scalability due to its modular design and integration with databases. It also operates with high efficiency in controlled environments with fixed seating arrangements.[6] However, the system does face certain challenges, including struggles with **real-time recognition** under varying environmental conditions, limitations in hardware and software that may hinder deployment in multi-classroom setups, and high implementation costs due to the advanced computational requirements of CNNs.[9] Despite these limitations, the proposed system offers a promising solution for **contactless attendance tracking**, particularly in environments where minimizing physical contact is crucial.

VII. Components

The system consists of several key components and actors, each playing a vital role in its functionality and operation. Students are the primary users who interact with the system to mark their attendance through voice recognition and provide vocal input for engagement analysis. Teachers are responsible for using the system to monitor and assess attendance, student engagement, and behavioral patterns over time, while also receiving alerts related to engagement issues or absenteeism. The LMS System is an external system integrated with the attendance system to facilitate the exchange of data, such as attendance records and engagement reports, ensuring seamless reporting and data management for educational institutions.

The use cases of the system cover several important functionalities: Automated Attendance, where students use voice recognition to mark their attendance automatically; Analyse Engagement, in which the system analyzes vocal expressions and patterns to assess students' engagement levels during class; Track Behavioural Patterns, allowing teachers to review and identify trends in students' behavior and engagement over time; Generate Alerts, notifying teachers about significant issues such as low engagement or persistent absenteeism; and Integrate with LMS, which ensures smooth data sharing between the system and the institution's Learning Management System for efficient reporting and academic tracking.

In terms of relationships, students directly interact with the Automated Attendance and Analyse Engagement features. Teachers, on the other hand, have access to all the functionalities of the system, including Track Behavioural Patterns and Generate Alerts, allowing them to effectively monitor student progress and engagement. The system continuously shares data with the LMS, ensuring real-time and seamless updates on attendance and engagement information for educators and administrators. This interconnected setup enhances the overall efficiency of tracking student participation and academic performance.

VIII. Working

The working of the system is designed to seamlessly handle attendance marking, engagement monitoring, and provide actionable insights for educators. The process begins with Student Interaction, where a student states their name or a designated phrase to mark their attendance. The system then verifies the voice using advanced voice recognition technology. Once the system has verified the student's voice, it confirms the attendance to the student in real-time.

Next, during the class, the Engagement Monitoring process kicks in. The system continually analyzes the student's voice for signs of engagement or emotional states such as attentiveness, confusion, or disinterest. This is done using advanced Natural Language Processing (NLP) and Machine Learning (ML) techniques, which assess vocal expressions and tone to determine the level of engagement. These insights are then provided to the educator, allowing them to monitor the class dynamically.

For Educator Interaction, educators access a user-friendly dashboard where they can view real-time attendance records as well as engagement metrics. The system also monitors the overall class engagement and issues alerts to educators when there are signs of low engagement or if students have been frequently absent. These alerts allow teachers to intervene promptly to address any concerns.

The system is integrated with an LMS (Learning Management System) to ensure that attendance records are automatically updated. After marking attendance, the system syncs the data with the LMS, and the LMS confirms that the update has been successfully made. This allows the attendance data to be centralized and easily accessed.

Finally, the data gathered through engagement analysis helps in Personalized Teaching. Educators use the insights about student engagement to adjust their teaching strategies, ensuring that they can cater to the needs of individual students, improving overall student outcomes. This personalized approach helps educators make data-driven decisions, improving both academic performance and student engagement.

IX. Algorithms

Voice Recognition is at the core of attendance marking. The system uses Speech-to-Text Models like Google's Speech-to-Text API or open-source alternatives such as DeepSpeech to convert spoken phrases or names into text.[8]This allows students to mark their attendance by simply stating their name or a designated phrase. To ensure accuracy and prevent impersonation, Speaker Identification and Verification algorithms, including MFCC (Mel Frequency Cepstral Coefficients), GMM (Gaussian Mixture Models), and DNN-based approaches, are used to verify the identity of the speaker.

Natural Language Processing (NLP) plays a significant role in Sentiment Analysis and Emotion Detection. The system uses Sentiment Analysis models, like BERT or DistilBERT, to classify the tone of the student's speech (e.g., engaged, bored, confused), which helps gauge their level of engagement. For more advanced analysis, Emotion Detection models are used to identify emotions based on vocal features and linguistic content, providing deeper insights into the emotional state of the student.

The system also includes Alert and Recommendation Systems that help educators take timely actions. Rule-Based Systems are used to trigger alerts for absenteeism or low engagement based on predefined thresholds. Additionally, Predictive Analytics models use historical data to predict which students might be at risk and recommend interventions tailored to individual needs.

X. Tools and Frameworks

Voice Recognition is supported by platforms like Google Cloud Speech-to-Text, IBM Watson Speech to Text, or Azure Speech Service, along with open-source solutions such as Kaldi, DeepSpeech, or Wav2Vec.[5] These tools enable high-accuracy transcription and speaker verification.

For Natural Language Processing, libraries such as NLTK, spaCy, and Hugging Face's Transformers are utilized for sentiment and emotion analysis. Custom-trained sentiment analysis models or Emotion Analysis APIs can be used to enhance emotion detection capabilities.

Machine Learning Frameworks like TensorFlow, PyTorch, Keras, and Scikit-learn are employed to build and train the models that analyze engagement levels, emotions, and predict student behavior based on past data.

To provide visual insights, Data Visualization and Dashboards are integrated using tools like Tableau, Power BI, Dash, D3.js, or Plotly. These tools allow educators to access interactive dashboards that display student engagement trends, attendance, and behavioral patterns over time.
Integration Tools such as REST APIs are used to connect the system with Learning Management Systems (LMS) like Moodle and Blackboard, facilitating seamless attendance and behavioral data sharing. For database management, SQL-based systems like MySQL and PostgreSQL, or NoSQL options like MongoDB and Firebase, store the data in a secure and organized manner.

Monitoring and Alerting is crucial for real-time engagement tracking. Prometheus with Grafana is used to monitor system performance, and notification systems like Twilio (for SMS/Email alerts) or the Slack API notify educators about critical issues such as low engagement or absenteeism, allowing for timely interventions.

XI. Conclusion

The Smart Attendance and Behavioural Analytic System functions through a seamless, automated process designed to improve both attendance tracking and student engagement analysis. When activated, the system begins by prompting students to either say their name or a specific phrase. Once the student speaks, the system uses Voice Recognition technology to listen and process the audio.
The system then performs a check If it successfully recognizes the student's voice, it marks them as present and proceeds to analyze their speech for emotional and engagement insights, such as their level of attentiveness or emotional state.
If the system fails to recognize the voice, it asks the student to repeat their phrase. If it still cannot identify the voice, the system logs the issue, ensuring that no attendance records are missed, and the teacher is alerted to the problem.Following the voice recognition process, the system carries out Engagement Analysis, where it examines various vocal cues to assess how engaged or attentive the student appears. These insights are crucial for understanding student behavior beyond just attendance.
Finally, the system updates the Dashboard for teachers, providing real-time feedback on attendance and engagement levels. If issues such as low engagement or absenteeism are detected, the system generates Alerts to notify educators, enabling them to take prompt action. This process not only streamlines attendance but also empowers educators with valuable insights to support and enhance student learning experiences.

XI. References

[1] A. Kumar and B. Singh, "IoT-Based Smart Attendance System Using Face Recognition and Cloud Storage," in Proc. IEEE Int. Conf. on Computation and Communication Technologies (ICCCT), 2022, pp. 112–118.
[2] P. Sharma, M. Gupta, and R. Verma, "Automated Attendance System Using Deep Learning-Based Face Recognition," in Proc. Int. Conf. on Machine Learning and Cybernetics (ICMLC), 2021, pp. 213–219.
[3] S. Raj, K. Kannan, and V. Kumar, "Voice and Image-Based Smart Attendance System Using AI," in Proc. IEEE Conf. on Artificial Intelligence and IoT (ICAIIOT), 2023, pp. 56–62.
[4] J. Park, H. Kim, and S. Choi, "Human Behavior and Attendance Tracking Using Deep Learning," in Proc. IEEE Conf. on Pattern Recognition and Computer Vision (PRCV), 2022, pp. 487–493.
[5] M. T. Islam, A. Rahman, and S. Ahmed, "A Hybrid Face and Voice-Based Smart Attendance System," in Proc. IEEE Int. Symp. on Smart Systems and Innovations (ISSSI), 2023, pp. 102–108.
[6] S. Gupta, R. Sharma, and P. Kumar, "Real-Time Face Recognition-Based Attendance System Using CNN and OpenCV," in Proc. IEEE Int. Conf. on Computational Intelligence and Data Science (ICCIDS), 2021, pp. 215–221.
[7] H. Lee, J. Kim, and K. Park, "Deep Learning-Based Smart Classroom Attendance Monitoring System," in Proc. IEEE Conf. on Internet of Things and AI Applications (IoTAIA), 2022, pp. 88–94.
[8] M. Patel, A. Shah, and P. Mehta, "Voice Recognition-Based Attendance System for Secure Authentication," in Proc. Int. Conf. on Speech Processing and AI (ICSPAI), 2023, pp. 312–318.
[9] T. Wang and Y. Zhang, "Behavioral Analytics in Smart Environments: A Deep Learning Approach," in Proc. IEEE Int. Conf. on Smart Computing (SMARTCOMP), 2022, pp. 134–140.
[10] S. Bose, K. Nair, and A. Iyer, "Hybrid Face and Voice Recognition-Based Attendance System Using AI," in Proc. IEEE Conf. on Computational Intelligence (CI), 2023, pp. 205–210.
[11] J. Park and S. Choi, "Human Activity Recognition for Smart Attendance Using CNN and LSTM," in Proc. IEEE Int. Symp. on Deep Learning and AI (DLAI), 2022, pp. 410–416.
[12] R. Kumar and V. Reddy, "IoT-Based Attendance Monitoring System Using Face and Voice Biometrics," in Proc. Int. Conf. on Emerging Trends in IoT and AI (ETIAI), 2023, pp. 89–96.
[13] A. Tan and L. Wei, "Deep Learning-Based Speaker Verification for Smart Attendance," in Proc. IEEE Conf. on Audio and Speech Processing (ICASSP), 2022, pp. 1453–1459.
[14] D. Singh, M. Sinha, and R. Gupta, "Smart Classroom Attendance System Using AI and Behavior Tracking," in Proc. IEEE Conf. on Smart Technologies in Education (STE), 2023, pp. 512–518.
[15] M. Jha and S. Roy, "Hybrid Face-Voice Authentication for Secure Attendance Systems," in Proc. IEEE Int. Conf. on Machine Learning and Cybersecurity (ICMLC), 2021, pp. 75–81.
[16] X. Liu and W. Chen, "Human Activity Recognition in Smart Environments Using CNN and Transformer

Models," in Proc. IEEE Conf. on Deep Learning and IoT (DL-IoT), 2022, pp. 390–397.
[17] K. Saito and Y. Nakamura, "AI-Driven Behavioral Analysis for Smart Attendance and Security Systems," in Proc. IEEE Conf. on Smart Computing (SMARTCOMP), 2022, pp. 103–109.
[18] A. Das and R. Bhatia, "Edge AI for Smart Attendance Systems Using Face and Speech Recognition," in Proc. Int. Symp. on Edge Computing and AI (ECAI), 2023, pp. 420–426.
[19] T. Nguyen and H. Le, "Deep Neural Networks for Voice-Based Authentication in Attendance Systems," in Proc. IEEE Conf. on Speech Processing and AI (SPAI), 2021, pp. 267–273.
[20] M. Al-Turki and S. Al-Qahtani, "Privacy-Preserving Smart Attendance Systems Using Blockchain and AI," in Proc. IEEE Int. Conf. on Smart Cities and Security (SCS), 2023, pp. 312–318.
[21] S. Arya, H. Mesariya, and V. Parekh, "Smart Attendance System Using CNN," arXiv preprint arXiv:2004.14289, 2020.
[22] B.-T. Nguyen-Tat, M.-Q. Bui, and V. M. Ngo, "Automating Attendance Management in Human Resources: A Design Science Approach Using Computer Vision and Facial Recognition," arXiv preprint arXiv:2405.12633, 2024.
[23] A. Rao, "AttenFace: A Real-Time Attendance System Using Face Recognition," arXiv preprint arXiv:2211.07582, 2022.
[24] T. Abderraouf, A. A. Wassim, and S. Larabi, "An Embedded Intelligent System for Attendance Monitoring," arXiv preprint arXiv:2406.13694, 2024.
[25] M. Pantić, "Machine Analysis of Facial Expressions for Human-Computer Interaction," in Proc. IEEE Int. Conf. on Multimedia and Expo (ICME), 2003, pp. 572–575.
[26] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," Int. J. Comput. Vision, vol. 57, no. 2, pp. 137–154, 2004.
[27] G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.
[28] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2014, pp. 1701–1708.
[29] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2015, pp. 815–823.
[30] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv preprint arXiv:1409.1556, 2014.
[31] C. Szegedy et al., "Going Deeper with Convolutions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2015, pp. 1–9.
[32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Proc. Advances in Neural Information Processing Systems (NIPS), 2012, pp. 1097–1105.
[33] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," in Proc. British Machine Vision Conference (BMVC), 2015, pp. 41.1–41.12.
[34] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Int. J. Comput. Vision, vol. 60, no. 2, pp. 91–110, 2004.
[34] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2005, pp. 886–893.
[35] L. van der Maaten and G. Hinton, "Visualizing Data Using t-SNE," J. Mach. Learn. Res., vol. 9, pp. 2579–2605, 2008.
[36] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," arXiv preprint arXiv:1412.6980, 2014.
[37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770–778.
[38] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.
[39] I. Goodfellow et al., "Generative Adversarial Nets," in Proc. Advances in Neural Information Processing Systems (NIPS), 2014, pp. 2672–2680.
[40] A. Graves, A.-r. Mohamed, and G. Hinton, "Speech Recognition with Deep Recurrent Neural Networks," in Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2013, pp. 6645–6649.
[41] A. Graves and N. Jaitly, "Towards End-to-End Speech Recognition with Recurrent Neural Networks," in Proc. Int. Conf. on Machine Learning (ICML), 2014, pp. 1764–1772.
[42] D. Amodei et al., "Deep Speech 2: End-to-End Speech Recognition in English and Mandarin," in Proc. Int. Conf. on Machine Learning (ICML), 2016, pp. 173–182.
[43] A. Senior et al., "Improving DNN Speaker Independence with I-Vector Inputs," in Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2014, pp. 225–229.
[44] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," IEEE Trans. Knowl. Data Eng., vol. 22, no. 10, pp. 1345–1359, 2010.
[45] M. D. Zeiler and R. Fergus, "Visualizing and Understanding Convolutional Networks," in Proc. European Conf. Comput. Vis. (ECCV), 2014, pp. 818–833.
[46] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proc. IEEE Conf. Comput.