Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Image Processing Based Attendance System with Dual-Stage LSTM and Machine Learning Models

Dr Bechoo Lal[1*], Shankar Prasad Mitra[2], Sonu Rana[3], Manoj Kumar Padhi[4], Abhishek Gupta[5], Shiv Kumar Reddy[6]

[1,6]Department of Computer Science & Engineering Koneru Lakshmaiah Education Foundation(KLEF) KL University , Guntur Dist-522302, Andhra Pradesh, India,

[2,3]Deperment of Computer Science and Engineering, Brainware University, Barasat, Kolkata, West Bengal, Pin 700125.

[4]Department of Computer Science and Engineering, Centurion University of Technology and Management, Paralakhemundi, Odisha, India

[5]Department of Computer Science and Engineering, Vishveshwarya Group of Institutions, Gautam Budh Nagar, UP, India

ARTICLEINFO

Received: 29 Dec 2024

Revised: 12 Feb 2025

Accepted: 27 Feb 2025

ABSTRACT

Introduction: The researchers introduced an Image Processing Based Attendance System with Dual-Stage LSTM and Machine Learning Models a pioneering solution in automated attendance tracking. This system innovatively combines the precision of facial landmark detection with the advanced learning capabilities of a Dual-Stage Long Short-Term Memory (LSTM) network.

Objectives:

The Advanced Attendance Management System with Dual-Stage LSTM and Facial Landmark Detection represents a comprehensive approach to automating attendance tracking through state-of-the-art facial recognition technologies. This research study delineates the system's architecture, detailing each module's function within the framework. The objective function used in this work is Sparse categorical Cross Entropy. The accuracy measure is calculated and reported at each epoch, and the weights are set using the efficient ADAM optimization technique.

Methods: This research study is designed to address the complexities of real-world environments, the system sets a new benchmark in recognizing and tracking individual faces over time. The Dual-Stage LSTM for face recognition operates by extracting and analyzing immediate facial features to identify distinct characteristics in the short term. Subsequently, the second stage processes these features over extended periods, learning and adapting to temporal variations in appearance, ensuring accurate long-term recognition despite changes in facial attributes.

Results: In implementing this system, special attention has been given to enhancing accuracy and reducing false positives, critical parameters in any attendance management application. By processing facial data through this dual-stage approach, the system demonstrates remarkable proficiency in handling variability in lighting, orientation, and background conditions.

Conclusions:

The Advanced Attendance Management System, leveraging Dual-Stage LSTM and Facial Landmark Detection, represents a breakthrough in automating attendance with cutting-edge facial recognition technology. Beginning with high-resolution image capture in the Input Module, the system meticulously processes these images through stages including pre-processing for image refinement, facial detection using advanced algorithms like the Haar Cascade Classifier and Viola-Jones, and precise feature extraction via the dlib library.

Keywords: Advanced Attendance Management System, Dual-Stage LSTM, Facial Landmark Detection, Real-Time Data Inputs, Adaptive Learning Algorithm, Real-World Environments.

INTRODUCTION

Maintaining classroom attendance management in the evolving educational technology landscape remains a critical, yet often cumbersome, task for educators worldwide. Traditional attendance tracking methods, such as manual roll calls or sign-in sheets, are time-consuming and prone to errors, leading to inaccuracies in attendance

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

records. These challenges highlight the need for a more efficient, accurate, and automated solution to manage classroom attendance. Integrating digital technologies into the classroom has opened new avenues for innovation in attendance management systems (AMS). A modern AMS can streamline the attendance process, reduce administrative workload, and enhance the accuracy of attendance records, thereby allowing educators to focus more on teaching and less on administrative tasks.

Artificial intelligence (AI) and machine learning advancements have paved the way for developing sophisticated AMS solutions tailored to the classroom environment. These systems leverage facial recognition technology and advanced algorithms to automate attendance tracking, offering a seamless, non-intrusive way to monitor student attendance. Using AI-powered facial recognition, such systems can quickly identify students as they enter the classroom, significantly reducing the time spent on manual attendance procedures. Moreover, integrating these technologies simplifies attendance tracking and introduces flexibility and security that traditional methods cannot offer. For instance, digital attendance systems can quickly adapt to various class sizes and settings, from small seminar rooms to large lecture halls, and ensure that attendance data is securely stored and managed.

The following approaches are used to recognize people: • Content-based approach • Luminance-based approach For face detection, ad a boost algorithm, float boost algorithm, neural network, a da boost algorithm, support vector machine, and Nave Baye's classifier have all been presented. With the aid of a rapid face detection algorithm, the performance of a facial recognition software technique for autonomous attendance for children in the school scenario without identification is improved. Using the proposed system, this method was used to recognize people in classroom pictures. There are two types of face recognition techniques: appearance-based and feature-based. Appearance-based employs the texture feature, which may be associated with a given head or specific areas. Feature-based utilizes physical features such as the lips, nose, eyes, eyebrows, cheeks, and their relationships. Face templates were created using analytical techniques such as Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA), and Neural Networks, as well as Eigen-faces [1-5]. The illumination invariant method reduces the illumination impact within the lecture.

The use rateof facial recognition software varies depending on the extraction technique of features employed. Face recognition algorithms are used with traditional algorithms, such as neural networks and fuzzy logic, to improve recognition accuracy [1]. Changes in head position and environmental light substantially impact the effectiveness of face recognition systems. Better attributes have been generated to solve this difficulty [27]. Many operations, such as object identification, tracking, and alignment, rely on landmark annotation. In their research, Riopka et al. [18], Cristinacce et al. [5], and Beumer et al. [3] found that landmarks are required for good face recognition [3]. Many research, such as face identification, facial expression recognition, and face tracking, rely heavily on face landmark detection [10].

However, variability in facial appearance due to factors such as lighting conditions, aging, and temporary changes (e.g., hairstyles, facial hair) poses significant challenges to maintaining high levels of accuracy over time. Our system employs a novel Dual-Stage Long Short-Term Memory (DS-LSTM) network architecture to overcome these obstacles. This architecture is specifically tailored to enhance the system's ability to recognize facial features in real time and adapt to their temporal variations, ensuring consistent and accurate identification across diverse conditions and over extended periods.

Furthermore, the system incorporates state-of-the-art facial landmark detection algorithms. These algorithms provide a granular analysis of facial features, allowing for precise identification even in scenarios of partial occlusion or varying facial expressions. By integrating these technologies, the proposed system offers a robust, adaptable, and user-friendly solution for attendance management that significantly outperforms traditional methods.

The rest of the paper is divided into different sections. Section 2 discusses a literature survey on face recognition and the importance of recurrent neural networks. Section 3 presents the proposed methodology, containing facial detection, facial landmark extraction, distance and slope measures, and DS-LSTM networks. Section 4 discusses about experimental results and discussions obtained. Finally, Section 5 concludes the paper by briefly describing the proposed methodology and plans for extending the work.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The quest for efficient and accurate attendance management systems (AMS) has led researchers to explore various biometric and image processing techniques. This section reviews the related works in the domain, highlighting diverse approaches and their outcomes, setting a context for our proposed "Advanced Attendance Management System with Dual-Stage LSTM and Facial Landmark Detection."

Studies have predominantly focused on applying facial recognition technologies, with works by Zhang et al. (2016) and Sharma et al. (2018) demonstrating the potential of utilizing facial detection and identification algorithms to streamline attendance processes in educational and corporate contexts. These initiatives, however, mainly employed single-layer neural network models, which, despite their effectiveness in controlled settings, face challenges with facial appearance variations over time.Sajid et al. [20] ventured into AMS using Retinal Detection, integrating Infrared (IR), RFID components, an LCD, and an Arduino UNO. Their system scans the lecture room through video pictures, displaying the computed number of attendees on an LCD. However, the system only achieved a 57 percent accuracy rate and was criticized for its time-consuming operation. In contrast, facial recognition technologies have shown promising improvements. An approach [23] employing a Multi-layer Neural Network for face recognition after initial detection reported an efficiency of 80 percent. Meanwhile, Tan. L et al. [11] leveraged the Gabor wavelet for face detection alongside K-Nearest Neighbor classification, achieving an 88 percent average precision and significantly advancing face recognition accuracy.

Samuel Lukas et al. [12] combined Discrete Wavelet Transform (DWT), Discrete Cosine Transforms (DCT), and Radial Basis Function Network (RBFN) in their Student Attendance system. Despite reaching an efficiency of 82 percent, their system sometimes confused students with others. Priyanka Wagh et al. [26] analyzed several facial recognition algorithms, including Principal Component Analysis (PCA), Eigenface, Support Vector Machines (SVM), and Neural Networks, to assess their success rates. Further enhancing identification processes, Abhishek Jha et al. [8] integrated statistical approaches PCA and LDA, collecting facial features from images for more precise human face identification.

Tata Sutabriet al. [23] proposed a CNN-based system for face detection and identification, utilizing adlib's CNN and deep metric modeling for embedding, with K-NN for identification, showcasing improvement in face attribute identification performance. Tan. L [11] improved face attribute recognition on the CelebA and LFWA databases by introducing a spatial transformer network, enhancing average performance by up to 4%.

Cherifi Dalila et al. [4] developed a Facial Recognition Software Registration System using both general (PCA, FLD, DCT, DWT) and regional (SIFT, LBP) methods, tested across various conditions, including posture, lighting, and facial expression changes. Nazare et al. [14] applied the Viola and Jones method for facial identification, achieving an 85 percent accuracy in detection but only a 53.1 percent success rate in recognition.

Papageorgiou et al. [16] suggested a continuous picture-taking AMS, identifying the most precise image for further analysis, a method that ensured varied gesture captures from multiple angles. Francisco et al. [17] explored a probabilistic approach with local binary patterns and weighted masks for facial feature extraction, focusing on crucial facial regions for identification. Stefano Arca et al. [2] employed Gabor Filters for facial calibration points identification, recommending an optimal 31 points for attendance registration with minimal computational demand.

OBJECTIVES

The objective of Long Short-Term Memory (LSTM) networks in processing temporal data has also been extensively explored. Graves et al. (2013) and Sutskever et al. (2014) have highlighted LSTMs' capacity for sequence prediction, laying the groundwork for their application in AMS to adapt.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

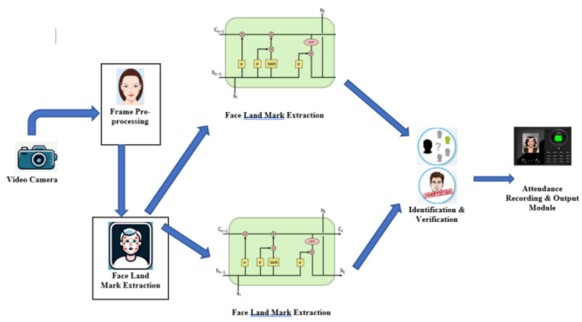

Figure 1. Overview of the Proposed System

to pattern changes. Nonetheless, the dual-stage LSTM networks' application—offering a nuanced, layered approach to learning temporal dependencies—remains largely uncharted in AMS, despite their success in fields like speech recognition, as evidenced by Chen et al. (2017). Similarly, advances in facial landmark detection by Kazemi et al. (2014) and Sun et al. (2013) have significantly improved facial feature analysis accuracy. However,integrating these methodologies with LSTM networks for enhanced facial recognition in attendance systems is a novel endeavor. Additionally, the critical aspect of privacy and security in AMS employing facial recognition has been addressed in recent literature, with Martinez-Diaz et al. (2019) discussing biometric systems' privacy challenges. Our proposed "Advanced Attendance Management System with Dual-Stage LSTM and Facial Landmark Detection" not only synthesizes these technological advancements but also innovates by combining dual-stage LSTM networks with facial landmark detection in a privacy-conscious framework, thereby setting a new standard for secure and efficient attendance management solutions. This integration uniquely addresses previous systems' limitations in managing temporal facial feature variations, marking a significant leap forward in the AMS domain.

METHODS

Input Module

The system initiates at the Input Module, capturing real-time video footage or images from the classroom or designated attendance area. This crucial entry point employs sophisticated cameras or video input devices to collect high-resolution visual data conducive to accurate facial detection. The module is the foundation upon which the subsequent processing stages are built, ensuring that initial data capture is efficient and of sufficient quality for further analysis.

Pre-processing Module

Following data capture, the pre-processing module undertakes the input images or video frame refinement. This stage is instrumental in enhancing the visual quality through resizing, normalization, and grayscale conversion processes. Such preliminary adjustments are vital for optimizing the images for facial detection, ensuring the system can accurately identify faces under varying lighting conditions and angles.

Facial Detection Module

At the core of the Advanced Attendance Management System lies the Facial Detection Module, which leverages the Haar Cascade Classifier and Viola-Jones algorithm for robust facial detection within the processed images. This module excels at rapidly identifying faces by scanning the visual data for Haar-like features and employing the integral image concept for swift feature detection. The efficacy and efficiency of the Viola-Jones algorithm are well-documented, making it a cornerstone for real-time face detection across various settings, including classroom environments. Its inclusion in the system significantly bolsters the system's capacity to accurately pinpoint faces, thereby setting the stage for meticulous feature extraction (Vikram & Padmavathi, 2017; Reddy & Nithya, 2023; Suryavanshi, Dubey, & Sharma, 2022). By harnessing the power of these algorithms, the system ensures reliable facial recognition, which is essential for the subsequent stages of attendance management.

Feature Extraction Module

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The Advanced Attendance Management System's Feature Extraction Module harnesses the dlib library's capabilities for precise facial feature extraction using detailed landmark points, making it a cornerstone of the system's facial recognition process. Dlib, renowned for its robust machine learning algorithms and versatility, excels in facial landmark detection by employing a pre-trained model that identifies 68 specific points on the face. This model delineates critical areas such as the jawline, eyebrows, eyes, nose, and mouth, facilitating the accurate mapping of facial features. The process initiates with dlib's efficient face detector that locates faces within an image, which the shape predictor then analyzes to pinpoint and map the landmark points. This meticulous extraction of facial landmarks enables the system to capture essential features like the distances between the eyes, the shape of the mouth, and the contour of the jawline, which is critical to the precise identification of individuals.
Integrating the dlib library into the Feature Extraction Module significantly enhances the Advanced Attendance Management System's capability to identify individuals accurately and efficiently, a critical requirement for effective attendance tracking. The choice of dlib offers numerous advantages, such as high accuracy in landmark detection and speed optimized for real-time processing, ensuring the system's adaptability to live classroom environments. Furthermore, dlib's training on diverse datasets guarantees the system's ability to recognize facial features across different demographics, enhancing its utility in varied settings. Through this integration, the system achieves high levels of precision in facial feature analysis. It ensures operational efficiency, making it a sophisticated solution for attendance management challenges in educational and corporate contexts.

Dual-Stage LSTM Network

The Dual-Stage LSTM (Long Short-Term Memory) Network within the Advanced Attendance Management System ingeniously capitalizes on the LSTM's inherent ability to process sequential data, tailoring it to the nuanced requirements of facial recognition for attendance management. This sophisticated architecture is engineered to discern and adapt to the temporal dynamics of facial features, ensuring high fidelity in attendance tracking over both short and long durations. The system seamlessly integrates this dual-stage approach to offer unparalleled precision in recognizing individuals under varying conditions and across time spans, accounting for daily fluctuations as well as more significant changes in appearance.

Stage 1: Short-term Memory Processing

The initial stage of the Dual-Stage LSTM Network focuses on short-term memory processing, employing the LSTM's capacity for handling sequences to identify individuals based on recent facial characteristics accurately. The LSTM unit, at itscore, follows the general equation:

$$h_t = O_t \odot \tanh(c_t) \text{.......(1)}$$

where h_t represents the output at time t, o_t is the output gate, $c_t$is the cell state, and $\odot$ denotes element-wise multiplication. For short-term processing, this stage leverages the LSTM's ability to remember and process information over short sequences, which is crucial for maintaining accuracy with daily variations in appearance.

Stage 2: Long-term Memory Processing

Progressing to the second stage, the network extends its prowess to long-term memory processing, encapsulating the LSTM's profound capability to learn from extended data sequences. This facilitates the system's adaptation to substantial changes in an individual's appearance over time. The long-term processing can be elucidated through the LSTM's cell state update equation:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \text{.....(2)}$$

Here, c_{t-1} is the previous cell state, f_t is the forget gate, i_t is the input gate, and $\tilde{c}_t$is the candidate cell state. This equation is pivotal for the system's ability to incorporate and remember significant alterations in facial features, ensuring consistent identification accuracy over prolonged periods.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Integrating the Dual-Stage LSTM Network into the Advanced Attendance Management System dramatically elevates its ability to track attendance through sophisticated facial recognition accurately. The first stage ensures adaptability to daily appearance variations, while the second stage guarantees sustained accuracy amidst more pronounced changes over time. This dual approach enhances the system's resilience to facial feature variability and assures a high degree of precision in attendance tracking, distinguishing it as a forward-thinking solution in automated attendance management.

The Dual-Stage LSTM Network's design, emphasizing sequential data processing and temporal adaptation, minimizes identification errors that could arise from conventional recognition methods. By learning the unique trajectory of everyone's facial features, the system delivers a personalized, accurate, and efficient attendance tracking solution, illustrating the potential of combining advanced neural network architectures with practical applications in attendance management.

Identification and Verification Module

The Identification and Verification Module is a key component of the Advanced Attendance Management System, tasked with matching processed facial features against a database of known individuals to accurately confirm their identity and attendance. This process is underpinned using mathematical and computational principles, particularly distance metrics or similarity scores, which quantify the closeness between the input feature vector and the stored profiles. A standard method involves calculating the Euclidean distance between two feature vectors and identifying the individual by selecting the database entry that yields the minimum distance, indicating the highest similarity.

For verification, the system employs a thresholding approach, where identity verification occurs if the similarity score surpasses a predefined threshold, ensuring accurate and reliable attendance marking. This mechanism is pivotal for enhancing the system's efficacy, leveraging machine learning models to improve identification accuracy continuously. By refining the process of distinguishing individuals and adapting to new data, the Identification and Verification Module significantly bolsters the system's precision and reliability, illustrating the integration of sophisticated computational techniques in practical applications for automated, secure attendance management.

Attendance Recording Module

The Attendance Recording Module is an essential component of the Advanced Attendance Management System, responsible for accurately logging the attendance of individuals once they have been identified and verified. This module updates the system's database to reflect the attendance status of an individual on a given day, incorporating timestamping to record the exact time of attendance. The process can be simplified into an operation where attendance A for individual i on day d is marked present (P), potentially with a timestamp t.

Beyond simple attendance logging, this module also facilitates the calculation of attendance statistics, such as the total number of days present, by aggregating daily attendance data. This functionality ensures the secure and efficient management of attendance records and supports the analysis of attendance trends and patterns over time.

Integrating this module significantly enhances the system's utility, offering a scalable and secure solution for managing attendance data. It enables easy access to reliable attendance records for administrative and educational analysis, highlighting the system's ability to leverage technology to improve the efficiency and effectiveness of attendance management.

Output Module

The Output Module is a critical component of the Advanced Attendance Management System, designed to aggregate, summarize, and visually present the attendance data to stakeholders like administrators, teachers, and students. This module transforms raw attendance records into comprehensive reports, real-time updates, and statistical analyses, facilitating easy access and interpretation of attendance information. Although it doesn't directly involve complex mathematical equations, it supports calculating metrics such as the average attendance rate, enhancing the system's analytical capabilities.

Integrating this module significantly improves the system's delivery of actionable insights and transparent communication regarding attendance trends and patterns. Customizing outputs according to specific needs ensures

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

that all users can effectively monitor, assess, and make informed decisions based on attendance data. Ultimately, the Output Module elevates the Advanced Attendance Management System by making attendance management more accessible, understandable, and actionable for improving overall attendance and engagement.

Experimental Results

This section outlines the experimental framework, emphasizing the dataset utilized and the implementation specifics. The experimental activities are conducted on a workstation equipped with an Intel Core™ i7-9750H CPU and 16.00 GB of RAM, running Ubuntu Linux, and leveraging Python 3.7 for development. For the initial phase of facial detection, we employ the robust dlib library, renowned for its accuracy in identifying facial features. Subsequent facial feature extraction leverages the advanced capabilities of the Dual-Stage LSTM Network, as delineated in our methodology section, to process and analyze temporal variations in facial features with unparalleled precision.

Continuing, we underscore the merits of our Advanced Attendance Management System by conducting a comparative analysis against prevailing models, particularly emphasizing its enhanced adaptability and efficiency in handling both short-term and long-term facial feature variations. A detailed evaluation of our system's performance is presented, benchmarking it against other state-of-the-art attendance management systems to highlight its superior accuracy and operational efficiency. Notably, our system's identification and verification module is based on a customized Dual-Stage LSTM model, which significantly deviates a significant deviation from traditional approaches and is central to our innovative strategy. As previously outlined, this comparative analysis is complemented by graphs and tables to represent our system's advantages. Through this comprehensive examination, we aim to substantiate the effectiveness and benefits of the Advanced Attendance Management System, setting a new benchmark in automated attendance tracking technologies.

Databases

The CMU-PIE database [30] contains over 40,000 images featuring 68 individuals. This extensive collection includes more than 600 images per subject, captured from 13 different angles (showcasing variations in head yaw and pitch), under 43 distinct lighting conditions, and depicting 4 facial expressions (neutral, talking, blinking, and smiling). To assess the performance of the proposed method for pose-invariant face recognition, we specifically selected images that feature ambient lighting, a neutral facial expression, and variations in the yaw angle, resulting in a dataset of 9 images per subject, totaling 612 images.

Two approaches are utilized to train the base learners for categorizing the pose of a facial image. The initial approach involves teaching the base learner set on semi-frontal images (those within a pose angle range of ±45 degrees), using exclusive images positioned at a 0-degree angle. The alternate approach trains the base learners on profile images (captured at pose angles of ±60 degrees and ±90 degrees), incorporating images at pose angles of −90 degrees and 60 degrees for this purpose. Consequently, for each individual featured in the CMU-PIE database, 3 images are utilized during the training phase of the base learners.

Comparison with Other Methods

ResNet50 [30]:

Cao. Q [30] introduced VGGFace2, a comprehensive facial recognition dataset,to enhance model performance across diverse poses and age groups. The dataset, featuring over 3 million images of more than 9000 subjects, is meticulously annotated to include a wide range of demographic characteristics, facial expressions, and environmental conditions, emphasizing the variability in pose and age. By employing advanced convolutional neural network (CNN) architectures for deep learning, the research team demonstrates the dataset's capability to improve the accuracy and robustness of to enhance facial recognition systems significantly.to improve the accuracy and robustness of facial recognition systems significantly.

MobileFacNet [31]:

W. Xie and A. Zisserman [30] explore the efficacy of multicolumn networks in enhancing facial recognition accuracy, specifically through the development of MobileFaceNet, which achieved a notable 79% accuracy. This approach leverages a multicolumn architecture that integrates multiple convolutional neural network (CNN)

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

columns, each designed to capture distinct facial features and characteristics, offering a comprehensive analysis from multiple perspectives.

LightCNN-V2 [32]:

Wu et al., [32] propose an innovative lightweight convolutional neural network (CNN) optimized for facial recognition tasks, even in datasets with noisy labels. Central to their approach is a streamlined CNN architecture that emphasizes computational efficiency without compromising on the accuracy of face representation. A key aspect of their methodology is a robust training strategy designed to mitigate the impact of inaccurately labeled data, incorporating noise reduction and label correction techniques to enhance training quality. Through depth-wise separable convolutions and efficiency-driven design choices, the network excels in extracting essential facial features, demonstrating its superiority in performance when benchmarked against traditional models on standard datasets. This study not only showcases the potential of lightweight CNNs in achieving high accuracy in facial recognition with minimal computational resources but also highlights the importance of addressing noisy labels to improve model reliability and applicability in resource-constrained scenarios.

Deepface[33]:

DeepFace [33], a face verification system that significantly reduces the accuracy gap between computer and human recognition. Utilizing a deep nine-layer neural network trained on a dataset of 4 million images from over 4,000 individuals, DeepFace leverages 3D modeling for facial alignment, standardizing images for improved consistency. This method allowed DeepFace to approach human-level performance on the Labeled Faces in the Wild (LFW) dataset, showcasing the potential of deep learning in achieving advanced facial verification accuracy.

VGGFace [34]:

O. M. Parkhi [34] introduced a novel deep-learning approach to face recognition using a specialized convolutional neural network (CNN) architecture. Training this deep CNN on a large-scale dataset is central to their methodology, enabling the model to learn from a wide variety of facial expressions, features, and individual variations. Through rigorous benchmarking against established face recognition benchmarks, the authors demonstrate the model's superior accuracy and explore optimization techniques to enhance its performance further. This study marks a significant advancement in face recognition technology, showcasing the deep learning model's ability to closely match human-level identification capabilities by effectively capturing and analyzing complex facial patterns.

RESULTS

A parameter derived from camera calibration [3] is used to correct camera distortion. The objective function used in this work is Sparse categorical Cross Entropy. The accuracy measure is calculated and reported at each epoch, and the weights are set using the efficient ADAM optimization technique. The LSTM will be trained for 1,000 epochs. The method, on the other hand, converges quickly, needing just 50 epochs. This is shown in fig.2 with Model Loss vs Epochs throughout the training phase. A fresh random input sequence will be generated for the network to fit on every epoch. Each epoch, the recommended model is run on the random sequences, and the log loss and classification accuracy are printed. This demonstrates how well the model has generalized the sequence classification issue solution. The Accuracy versus Epoch plot is shown in Figure 3, to provide a clear understanding of the generalized solution.

cThe confusion matrix analysis (Table 1) for face pose classification with $N_{pose} = 5$ on the CMU-PIE dataset reveals how the classifier performs across different face pose angles. $Class\,1$, representing the profile view a -90^{o}, achieves the highest metrics with an accuracy of 92.8%, recall of 94%, and a precision of 93%, culminating in an F-1 score of 91%. This indicates that the classifier excels at identifying faces at this extreme profile angle.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

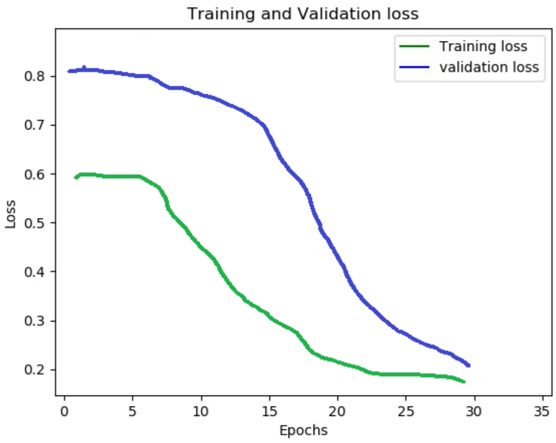

Figure 2. Model Loss vs Epochs

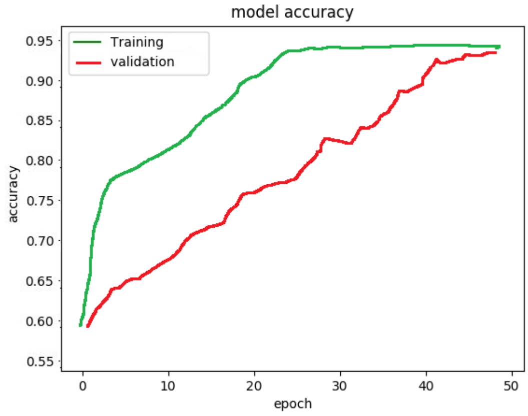

Figure 3. Model Accuracy vs Epochs

2, with the -60^o pose, shows a slight dip in performance with an accuracy of 90% and an F-1 score of 89%, suggesting a high level of effectiveness in recognizing this pose as well. The semi-frontal poses encompassed in Class 3, at -45^o, and $+45^o$present more of a challenge to the classifier, as accuracy falls to 85%,. However, recall and precision remain relatively high at 88% and 87%, respectively, resulting in an F-1 score of 88%. Near-profile and full-profile poses, represented by Classes 4 and 5 at $+60^o$ and $+90^o$, respectively, both register an accuracy of 84%. While their recall and precision figures are similar, Class 4 has a marginally higher recall and Class 5 a slightly greater precision, with F-1 scores indicating a robust performance of 88% and 86%. Overall, the classifier demonstrates a solid capability to distinguish between varying face poses, with its performance being particularly noteworthy at more extreme profile angles, as evidenced by the consistently high F-1 scores across all classes.

Table.1 Confusion matrix analysis results for face pose classification with Npose = 5 classes On

Class	**Angle Range**	**Accuracy**	**Recall**	**Precision**	**F-1 Score**
1	-90^o	0.928	0.94	0.93	0.91
2.	-60^o	0.90	0.89	0.90	0.89
3	$-45^o, +45^o$	0.85	0.88	0.87	0.88
4	$+60^o$	0.84	0.85	0.86	0.88
5	$+90^o$	0.84	0.82	0.84	0.86

Table 1 outlines a succinct comparison of face recognition methods based on their performance on the CMU-PIE dataset, ranked by accuracy (ACC) and Area Under the Curve (AUC) metrics. Beginning with ResNet50 [30], we see a competent performance with an accuracy of 77.8% and an AUC of 76.7%, which serves as a baseline for subsequent models. MobileFaceNet [31] improves upon this with a 79.6% accuracy and a 78.9% AUC, indicating enhanced recognition capabilities. Progressing to LightCNN-V2 [32], there is a notable jump in performance, achieving an 84.5% accuracy and an 85.4% AUC, which suggests its strength in differentiating between positive and negative classes.

DeepFace [33] advances the benchmarks significantly, attaining an 89.9% accuracy and a 90.2% AUC, reflecting the sophistication of its deep learning architecture. VGGFace [34] slightly eclipses DeepFace, posting an accuracy of 91.4% and an AUC of 90.8%, likely due to its deep convolutional networks trained on an extensive facial dataset. The DS-LSTM model, the focus of the proposed work, outperforms all other methods with the highest accuracy of 92.6% and an AUC of 92.4%, suggesting that its innovative use of a dual-stage LSTM network is particularly effective in capturing temporal variations in facial features, leading to robust face recognition performance. The DS-LSTM's leading results indicate its potential as a superior approach for face recognition tasks, especially where adaptability to facial feature changes is critical.

The ROC curve data (as shown in Figure 4) provided for the CMU-PIE dataset showcases the performance of various facial recognition models, including the proposed DS-LSTM model, regarding their False Positive Rate (FPR). The FPR is a critical measure in classification tasks, indicating the proportion of non-face instances mistakenly identified as faces by the model. Lower FPR values indicate a model's efficacyin accurately distinguishing faces from non-faces, a desirable trait for any facial recognition system. LightCNN-V2 demonstrates exceptional performance at lower FPR thresholds, suggesting robust discrimination capabilities. However, as the threshold for classification is relaxed, the DS-LSTM model exhibits a higher FPR than its counterparts, culminating in an FPR of 0.92 when all false positives are accepted.

This suggests that while the DS-LSTM model is potentially more inclusive, capturing a higher number of true positives, it also incurs more false positives across the board. Such a characteristic could be advantageous in applications where missing a true positive is particularly costly, but it may also result in an increased burden of verification checks due to the higher number of false alarms. Therefore, while the DS-LSTM model shows promise, its application would be best suited to environments where the cost of missing a true positive is greater than the cost of handling false positives.

Table 2. Accuracy (ACC) and Area Under the Curve (AUC) Comparison On CMU-PIE

S. No	**Method**	**ACC (%)**	**AUC (%)**
1	ResNet50 [29]	77.8	76.7
2.	MobileFaceNet [30]	79.6	78.9
3	LightCNN-V2 [31]	84.5	85.4
4	DeepFace [32]	89.9	90.2
5	VGGFace [33]	91.4	90.8
6	DS-LSTM (Proposed Work)	92.6	92.4

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

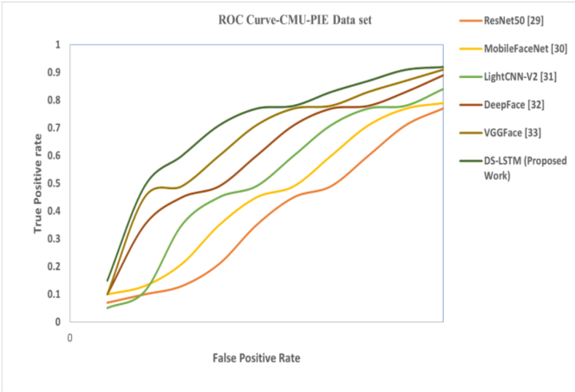

Figure 4. Comparison of ROC Curve.

(a)

(b)

(c)

(d)

Figure.5- Input/Output Images after processing different stages-
(a) Input Image, (b) Color to Grey Scale Converted Image,
(c) Normalized Image, (d) Landmark Detected Image.

We tested our methodology in a classroom photograph taken with a Logitech camera. The people are seated in front of the camera at a distance of 100 cm. But for the sake of illustration, we took a sample image from internet and applied our methodology. Figure 5-(a-d) depicts the steps involved in capturing a sample classroom input image, converting it to greyscale, using Histogram Normalization to correct contrast, Haar Cascade Face detection, and finally, Landmark recognition.

Figure 5 showcases the meticulous process of image transformation within a facial recognition system, culminating in a sophisticated analysis that identifies vital landmarks and achieves facial recognition. Beginning with the (a) Input Image, this initial stage captures the raw, unaltered photograph, full of the rich details and nuances inherent to any natural setting. The transformation journey continues with the conversion of this image into (b) Grayscale. This step eliminates the color data to emphasize the light intensity across the image, streamlining the complexity and focusing on the structural details crucial for further processing. Subsequently, the image undergoes (c)

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Normalization, aiming to standardize the pixel intensity levels, enhancing the overall image contrast, and ensuring uniformity in lighting conditions across different images. This is vital for consistent feature analysis.

The final transformation is depicted in (d) Landmark Detection and Facial Recognition, where the processed image is not only annotated with crucial facial landmarks—such as the eyes, nose, mouth, and jawline—but also subjected to facial recognition algorithms. This advanced stage involves identifying and marking key points on the face to accurately analyze distinct facial features and expressions. Moreover, it extends beyond mere landmark detection by integrating facial recognition capabilities, thereby matching the detected facial features against a database of known faces to establish identity. This comprehensive approach encapsulates the entire facial recognition process, from initial image capture to the final identification, highlighting the system's capability to recognize individuals based on their facial features accurately. Through these successive stages, Figure 5 effectively demonstrates the complex and nuanced process of preparing an image for facial recognition, showcasing the evolution from a simple photograph to an analyzed image ready for identity verification.

DISCUSSION

The Advanced Attendance Management System, leveraging Dual-Stage LSTM and Facial Landmark Detection, represents a breakthrough in automating attendance with cutting-edge facial recognition technology. Beginning with high-resolution image capture in the Input Module, the system meticulously processes these images through stages including pre-processing for image refinement, facial detection using advanced algorithms like the Haar Cascade Classifier and Viola-Jones, and precise feature extraction via the dlib library. The core innovation, the Dual-Stage LSTM Network, adeptly handles both short-term and long-term variations in facial features, significantly enhancing the system's recognition accuracy. Complemented by the Identification and Verification Module, the system ensures reliable identity confirmation, while the Attendance Recording Module efficiently logs and analyzes attendance data, all presented through the Output Module for easy stakeholder access.

Experimental evaluations and comparisons with existing models underscore the proposed system's superior performance, particularly in accuracy and Area Under the Curve (AUC) metrics, outperforming renowned methods like ResNet50, MobileFaceNet, and VGGFace. Such robust performance is further highlighted through confusion matrix analysis across various facial poses, demonstrating the system's adeptness at recognizing faces under diverse conditions. The transformation of images through different processing stages, culminating in landmark detection and facial recognition, showcases the system's comprehensive approach to preparing images for precise recognition tasks. Overall, this system sets a new standard in attendance management technologies, merging advanced neural network architectures with practical applications to offer a sophisticated, efficient solution for educational and corporate environments.

REFRENCES

[1] Anil.J , and Padma Suresh. L, "Literature survey on face and face expression recognition," 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT), 2016, pp. 1-6.
[2] Arca, Stefano, Paola Campadelli, and Raffaella Lanzarotti. "A face recognition system based on automatically determined facial fiducial points", Vol. 39, Iss. 3, 2006, pp. 432-443.
[3] Beumer. G, Bazen. A, and Veldhuis.R, "A landmark paper in face recognition," 7th International Conference on Automatic Face and Gesture Recognition (FGR06), 2006, pp. 66-78.
[4] Cherifi Dalila, CherfAouiFateh,Yacini Si Nabil and Nait-ali,"Fusion of Face Recognition Methods at Score level" , Proc. Of the IEEE International Conference on Bio-engineering for Smart Technologies, BioSMART 2016, Dubai, United Arab Emirates. pp.1-5.
[5] Cristinacce D, Cootes T, and Scott I, "A multi-stage approach to facial feature detection," In 15th British Machine Vision Conference, London, England, 2004, vol. 1, pp. 277–286.
[6] Emami, "Face Recognition using Eigenface or Fisherface", in Mastering OpenCV with Practical Computer Vision Projects. Pact Publishing, 2012. ISBN 1849517827.
[7] Islam S, Mahmud. A, Papeya A, Onny I and Uddin.J, "A Combined Feature Extraction Method for Automated Face Recognition in Classroom Environment," 3rd International Symposium on Signal Processing and Intelligent Recognition Systems, SIRS'17, Karnataka, India, 2017, vol 678, pp. 417-426.

[8] Jha. A, "Class Room Attendance System Using Facial Recognition System.", in International journal of Mathematical science technology and management,2007, Vol. 2, Iss.3, 2007.
[9] Kadry. S and Smaili. K, “A design and implementation of a wireless iris recognition attendance management system”, Information Technology and control, 2017, vol. 36, no. 3, pp. 323– 329.
[10] Liang. A, Liu. W, Li. L, Farid. M. R, and Le.V, “Accurate facial landmarks detection for frontal faces with extended tree structured models,” In Pattern Recognition (ICPR), 22nd International Conference on IEEE, 2014, pp. 538-543.
[11] Tan. L, Li. Z and Yu. Q, "Deep face attributes recognition using spatial transformer network," 2016 IEEE International Conference on Information and Automation (ICIA), 2016, pp. 1928-1932.
[12] Lukas. S, Mitra. A. R, Desanti. R. I and Krisnadi. D, "Student attendance system in classroom using face recognition technique," 2016 International Conference on Information and Communication Technology Convergence (ICTC), 2016, pp. 1032-1035.
[13] Ming Du, Aswin C. Sankaranarayanan, and Rama Chellappa ,"Robust Face Recognition From Multi-View Videos," in IEEE Transactions on Image Processing, 2014, vol. 23, no. 3, pp. 1105-1117.
[14] Nazare Kanchan Jayant, Surekha Borra "Attendance management system using hybrid face recognition techniques," 2016 Conference on Advances in Signal Processing (CASP), 2016, pp. 412-417.
[15] Nita Thakare, et al, “Face Detection And Recognition For Automatic Attendance System” by International Journal of technology and MobileComputing, 2016, Vol.5 Issue.4, pp. 74 – 78.
[16] Papageorgiou.C. P. and Oren. M, “A general framework for object detection,” Computer Vision, IEEE Int. Conf., January, 1998, pp.555-562.
[17] Pujol, Francisco A., and Juan Carlos García. "Computing the Principal Local Binary Patterns for face recognition using data mining tools." Expert Systems with Applications2012, Vol. 39, Iss.8,pp: 7165- 7172.
[18] Riopka. T, and Boult. T, “The eyes have it,” In Proceedings of ACM SIGMM Multimedia BiometricsMethods and Applications Workshop, Berkeley, CA, 2003, pp. 9–16.
[19] Riha. Z, Matyas V, Histogram and face recognition.2016, Faculty of Informatics Masaryk University.
[20] SajidM, R. Hussain and M. Usman, "A conceptual model for automated attendance marking system using facial recognition", Ninth International Conference on Digital Information Management (ICDIM 2014), 2014, pp. 7-10.
[21] ShubhiShriwastav and Dinesh Chandra Jain, “A Review on Face Recognition Attendance System”, The International Journal of Computer A publications, 2016, Vol. 143, Iss.8, pp.19-22.
[22] Shutao. L, Gong. D, and Yuan. Y, "Face recognition using Weber local descriptors," Neurocomputing, 2013,vol. 122, pp. 272– 283.
[23] Tata Sutabri, Pamungkur, Ade Kurniawan, and Raymond ErzSaragih, “Automatic Attendance System for University Student Using Face Recognition Based on Deep Learning”, International Journal of Machine Learning and Computing, 2019, Vol. 9, No. 5, pp. 668-674.
[24] Tripathi K. P, “A Comparative Study of Biometric technologies with relation to Human Interface”, in International Journal of Computer Applications,2018, Vol.14, No. 5, pp. 11-15.
[25] Tong Y, Liu X, Wheeler F. W, and Tu P. H, “Semi supervised facial landmark annotation,” Computer Vision and Image Understanding, 2012, Vol. 116, Iss. 8, pp. 922-935.
[26] Wagh P, Patil S, Chaudhari J and Thakare. R, "Attendance System based on Face Recognition using Eigen face and PCA Algorithms," 2015 International Conference on Green Computing and Internet of Things (ICGCIoT), 2015, pp. 303-308.
[27] Zhang Z, Wang. L, Zhu. Q, Chen. S.K, and Chen. Y, “Pose invariant face recognition using facial landmarks and Weber local descriptor,” Knowledge-Based Systems,2015, vol. 84, pp. 78– 88.
[28] SimT, Baker. S and Bsat. M, "The CMU pose illumination and expression database", IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 12, pp. 1615-1618, Dec. 2003.
[29] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, “VGGFace2:A dataset for recognising faces across pose and age,” in Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG), May 2018, pp. 67–74.
[30] W. Xie and A. Zisserman, “Multicolumn networks for face recognition,” 2018, arXiv:1807.09192.
[31] X. Wu, R. He, Z. Sun, and T. Tan, “A light CNN for deep face representation with noisy labels,” IEEE Trans. Inf. Forensics Security, vol. 13, no. 11, pp. 2884–2896, Nov. 2018.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[32] 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., Jun. 2014, pp. 1701–1708.

[33] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in Proc. British Mach. Vis. Conf. (BMVC), Swansea, U.K.: British Machine Vision Association, 2015, pp. 41.1–41.12