J-INTECH (Journal of Information and Technology)
Accredited Sinta 4 Ministry of Higher Education, Science and Technology
Republic of Indonesia SK No. 10/C/C3/DT.05.00/2025
E-ISSN: 2580-720X || P-ISSN: 2303-1425

Preventive Attendance Record using Photo from Mobile Phone and Printed Paper using CNN

Bradika Almandin Wisesa[1], Vivin Mahat Putri[2*], Evvin Faristasari[3], Sirlus Andreanto Jasman Duli[4], Silvia Agustin

[1,2,5]*Politeknik Manufaktur Negeri Bangka Belitung, Department of Informatics and Business, Jalan Raya Timah, Indonesia*
[3,4]*Politeknik Manufaktur Negeri Bangka Belitung, Department of Electronics and Agriculture, Jalan Raya Timah, Indonesia*

Keywords
CNN; Face Recognition; Prevention; System

****Corresponding Author:***
Bradika@polman-babel.ac.id

Abstract
Face-based attendance systems are increasingly popular for their ease of use, but they are susceptible to fraud, such as using photos or videos for unauthorized attendance. This study introduces a digital attendance system that combines facial recognition with liveness detection powered by Convolutional Neural Networks (CNN). Liveness verification is achieved by analyzing subtle movements and responses to ambient lighting. The dataset includes 30 facial images, encompassing both authentic and fraudulent samples. Testing demonstrates a facial recognition accuracy of 91.3% and effective spoofing detection in static and dynamic settings. This system provides a secure, fraud-resistant attendance solution ideal for educational and corporate settings. Further enhancements are suggested to improve performance across diverse facial expressions and lighting conditions.

1. Introduction

The advancement of digital technology has transformed attendance recording, with photo-based systems powered by face recognition gaining popularity in educational and professional settings for their efficiency and accuracy [1][2]. However, these systems are vulnerable to fraud, particularly proxy attendance, where static images, screenshots, or videos are used to deceive the system [2][3]. This security challenge underscores the need for robust liveness detection to verify the physical presence of individuals, ensuring tamper-resistant attendance tracking [2][3]. Previous studies have explored Convolutional Neural Networks (CNNs) for face recognition and liveness detection, but few have focused on preventing fraudulent attendance in software-based systems [4][3]. For instance, Wei et al. proposed an improved CNN model for liveness detection, achieving robust performance in distinguishing genuine faces from spoofed images, though primarily for security applications rather than attendance systems [3]. Elmahmudi and Ugail developed a deep face recognition system handling imperfect facial data, but it lacked specific anti-spoofing mechanisms for real-time attendance scenarios [4]. Romeo et al. incorporated motion detection in video-based analysis, relevant for liveness verification, but their focus was on mobility monitoring, not attendance tracking [1][5]. Mahum et al. explored deep neural networks for video surveillance, analyzing dynamic facial cues that could enhance liveness detection, yet their work targeted surveillance rather than attendance [6]. In the context of e-learning, Maisha and Shetu highlighted the potential of CNN-based systems but emphasized the need for robust anti-spoofing measures [7][8]. Other studies, such as those by Pandit et al. and Qin et al., utilized Multi-Task Cascaded Convolutional Neural Networks (MTCNN) for face detection, but they did not address spoofing in attendance systems [9][10][11]. Local efforts, such as Pratama and Santoso, implemented CNN-based face recognition for attendance but lacked liveness detection, leaving systems susceptible to fraud [12]. Unlike prior work, this research uniquely integrates CNN-based facial recognition with software-based liveness detection to address proxy attendance fraud, offering a hardware-independent solution tailored for educational and corporate environments [2][13]. By leveraging micro-movements (e.g., eye blinks, head tilts), skin texture, and lighting patterns, the system enhances security without requiring specialized equipment, distinguishing it from hardware-dependent approaches [2] . This novelty lies in combining these techniques into a practical, scalable attendance system, addressing gaps in real-time spoofing detection identified in previous studies [14][15]. The objective of this study is to develop a secure, software-based digital attendance system that integrates CNN-based facial recognition with liveness detection to prevent proxy attendance fraud using static images or screenshots, suitable for educational and corporate environments [2][14][13].

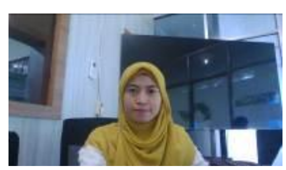
2. Research Method

This research employs an experimental quantitative method to design and evaluate a digital attendance system integrating facial recognition with liveness detection capabilities [14][16][17]. The main goal is to address vulnerabilities in photo-based attendance systems, particularly the risk of fraud using static images or facial screenshots [2][3]. The study involves data collection, image preprocessing, deep learning model training, and performance assessment [14]. The CNN model was trained on a system equipped with an Intel Core i7-10700 CPU, 16 GB RAM, and an NVIDIA GeForce RTX 3060 GPU . Training utilized Python 3.8, TensorFlow 2.6, and Keras libraries. Image preprocessing and optical flow analysis were performed using the OpenCV library [3][18]. The training process lasted around 12 hours for 100 epochs, with a batch size of 32.

This study involved active students from an Indonesian university who had prior experience with an online attendance system [8]. Fifty students voluntarily participated and consented to the use of their facial data [8]. The collected data was divided into two categories: valid attendance (authentic facial presence captured directly by the camera) and invalid attendance (faces presented via static images, screenshots, or other device displays to mimic fraudulent attempts) [2][3].

The primary data in this study include facial images captured using a laptop camera or webcam, liveness data derived from short videos and spoofing simulations (such as photos or images), and manually labeled data categorized as real or fake [2][3]. The dataset comprises 800 facial images, allocated as 70% for training, 15% for validation, and 15% for testing [14].

Table 1. Data Collection

Subjects	Picture
Sesa	
Evvin	
Sirlus	

Following the data collection process, Figure 1 illustrates the flowchart of the proposed digital attendance system, outlining the sequential steps from input capture to classification output [14]. The flowchart begins with facial image acquisition via a standard webcam, followed by preprocessing stages such as face cropping, grayscale conversion, and resizing to 224x224 pixels . It then depicts the CNN model training phase, where the preprocessed images are used to extract facial features and train the model for both recognition and liveness detection [14]. The liveness detection component, highlighted in the flowchart, involves analyzing micro-movements (e.g., eye blinks, head tilts) using optical flow, skin texture via Local Binary Patterns, and lighting distribution through pixel intensity gradients [3]. The final stage shows the system's binary classification output (live or spoof), which determines whether the input is a genuine face or a fraudulent attempt, such as a printed photo or screen-displayed image [2]. This visual representation clarifies the system's workflow, emphasizing its integration of facial recognition and liveness detection to prevent proxy attendance fraud, as discussed in [14]. These are flows of the system.

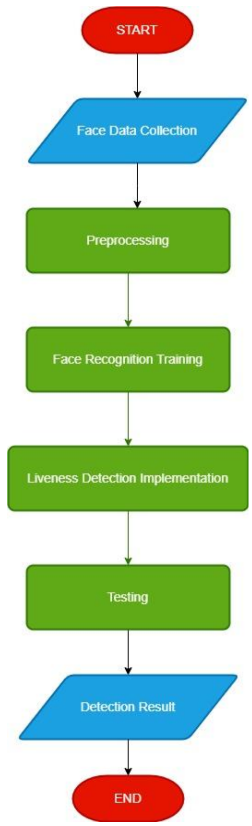


Figure 1. Flowchart of The System

Following the data collection process, Figure 1 illustrates the flowchart of the proposed digital attendance system, outlining the sequential steps from input capture to classification output [14]. The flowchart begins with facial image acquisition via a standard webcam, followed by preprocessing stages such as face cropping, grayscale conversion, and resizing to 224x224 pixels . It then depicts the CNN model training phase, where the preprocessed images are used to extract facial features and train the model for both recognition and liveness detection [14]. The liveness detection component, highlighted in the flowchart, involves analyzing micro-movements (e.g., eye blinks, head tilts) using optical flow, skin texture via Local Binary Patterns, and lighting distribution through pixel intensity gradients [3]. The final stage shows the system's binary classification output (live or spoof), which determines whether the input is a genuine face or a fraudulent attempt, such as a printed photo or screen-displayed image [2]. This visual representation clarifies the system's workflow, emphasizing its integration of facial recognition and liveness detection to prevent proxy attendance fraud, as discussed in [14].

The research followed these steps: (1) Data collection: facial images were captured using a standard webcam, and spoofing data was generated using printed photos, screenshots, and screen-displayed images to simulate fraudulent attempts [2][3]. (2) Image preprocessing: images were cropped to isolate faces, converted to grayscale, resized to 224x224 pixels, and augmented with rotations and lighting adjustments to enhance model robustness . (3) CNN model training: a CNN model was trained with multiple convolutional layers, max pooling, flattening, and dense layers using ReLU and softmax activations to extract facial features and detect liveness [14]. (4) Liveness detection: analyzed micro-movements (e.g., eye blinks, head tilts) using optical flow techniques, alongside skin texture and lighting distribution, to differentiate live faces from static images [3]. (5) System testing: evaluated using new data, with performance metrics including accuracy, precision, recall, and F1-score [14].

Convolutional Neural Network (CNN) merupakan metode yang umum digunakan dalam tugas-tugas penglihatan komputer, terutama dalam deteksi wajah karena kemampuannya dalam mengekstraksi dan mempelajari fitur visual secara hierarkis secara otomatis dari citra [14]. Struktur dasar CNN terdiri dari beberapa lapisan konvolusi yang bertugas mengenali pola visual lokal seperti tepi dan tekstur, diikuti oleh fungsi aktivasi seperti ReLU untuk menambahkan sifat non-linier, serta lapisan pooling yang bertujuan mengurangi dimensi spasial sambil tetap mempertahankan informasi penting [14]. Dalam sistem Pencatatan Kehadiran Preventif, CNN dimanfaatkan untuk mendeteksi dan mengekstraksi ciri khas wajah dari foto yang diambil melalui ponsel atau hasil pemindaian dokumen . Arsitektur CNN umumnya terdiri dari beberapa blok konvolusi dan pooling yang diakhiri dengan lapisan fully connected untuk menghasilkan representasi fitur wajah (embedding) atau melakukan klasifikasi [14]. Berkat kemampuannya yang tahan terhadap variasi pencahayaan, posisi, dan ekspresi wajah, CNN menjadi fondasi yang andal untuk membangun sistem kehadiran berbasis wajah yang akurat dan adaptif [4][14].

$$y_{i,j}^{(k)} = f\left(\sum_{m=1}^{M} \sum_{i=0}^{N} x_{i+m,j+n} \cdot w_{m,n}^{(k)} + b^{(k)}\right)$$

Table 2. Remarks of Each Symbol

Symbol	Value
$y_{i,j}^{(k)}$	output of the k-th feature map
$x_{i+m,j+n}$	output of the k-th feature map
$w_{m,n}^{(k)}$	weight of the k-th kernel
$b^{(k)}$	k-th bias
f	activation function (ReLU)

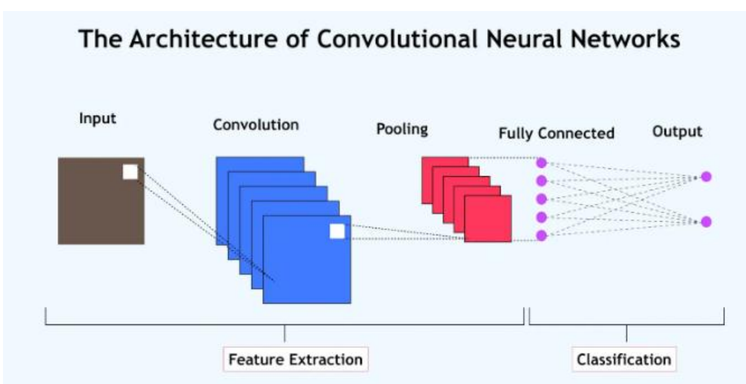


Figure 2. CNN Architecture

The core function of a Convolutional Neural Network (CNN) is comparable to solving a jigsaw puzzle, where the network identifies small elements like edges or shapes in an image and integrates them through deeper layers to create a comprehensive interpretation, resulting in classification or prediction [14]. CNNs are essential for streamlining and speeding up the learning process from visual data, with uses in facial recognition, medical imaging, autonomous vehicles, and natural language processing [19][20]. The adaptability of CNNs makes them suitable for diverse industries and applications [14]. The liveness detection system distinguishes live faces from spoofed images by evaluating visual indicators such as micro-movements, skin texture, and natural lighting patterns through techniques like optical flow, Local Binary Patterns, and pixel intensity gradient analysis [3].

Micro-movements, such as eye blinks, subtle head tilts, or slight facial muscle twitches, are critical indicators of liveness that static images cannot replicate. In this system, micro-movements are detected using optical flow analysis**,** specifically the Dense Optical Flow method based on the Farnebäck algorithm, implemented via the OpenCV library in Python. Optical flow refers to the pattern of apparent motion of objects within a sequence of images or video frames, caused by the relative motion between the camera and the scene.

The system captures a short video sequence (typically 2–3 seconds) or a series of consecutive frames from the camera input. The Farnebäck algorithm computes dense optical flow by estimating motion vectors for every pixel between pairs of consecutive frames. This generates a flow field that captures the direction and magnitude of pixel displacements.

The system focuses on regions of interest (ROIs) such as the eyes and mouth, where micro-movements are most pronounced. For example, an eye blink produces a distinct motion pattern in the flow field, characterized by rapid vertical displacement in the eyelid region. Similarly, subtle head tilts result in consistent horizontal or vertical flow vectors across the face. These motion patterns are quantified by calculating the average magnitude and direction of flow vectors within the ROIs.

A threshold is applied to the magnitude of the flow vectors to distinguish genuine micro-movements from noise or minor variations (e.g., camera shake). A minimum flow magnitude (e.g., 0.5 pixels per frame) is set to confirm liveness, ensuring that static images, which produce negligible flow, are classified as spoofed. Optical flow is robust to lighting variations and does not require specialized hardware, making it suitable for standard cameras. It effectively captures dynamic behaviors that are difficult to replicate in spoofing attempts. Optical flow can be sensitive to significant camera motion or background movement, which may introduce false positives. To mitigate this, the system normalizes the flow field by subtracting global motion estimates (e.g., using affine transformation) to isolate facial micro-movements.

Frame differences, which involve subtracting pixel intensities between consecutive frames to detect changes, were considered but not used as the primary method. Frame differences are simpler but less precise for capturing subtle motions, as they do not provide directional information and are highly sensitive to lighting changes or noise. Optical flow was chosen for its ability to model both magnitude and direction of motion, offering better discrimination of micro-movements. In addition to optical flow, the system employs statistical texture analysis using Local Binary Patterns (LBP) to evaluate skin texture and further distinguish live faces from static images. Printed photos or screen-displayed images often exhibit unnatural texture patterns, such as pixelation, moiré patterns, or lack of fine skin details.

The facial image is preprocessed by cropping the face region and converting it to grayscale. LBP is computed by comparing each pixel with its neighbors (typically in a 3x3 or 5x5 neighborhood) to generate a binary code, which is then aggregated into a histogram. This histogram captures the distribution of texture patterns, such as edges, smoothness, or irregularities. Live faces typically exhibit complex, fine-grained texture patterns due to natural skin features (e.g., pores, wrinkles). In contrast, printed photos or screen images show artifacts like uniform smoothness or repetitive patterns. The system uses a pretrained classifier (e.g., a Support Vector Machine) to compare the LBP histogram against a dataset of live and spoofed samples. The LBP analysis is performed using OpenCV's implementation of uniform LBP, with a radius of 1 and 8 neighboring points, resulting in a 59-bin histogram. The classifier is trained on the 800-image dataset, with 70% used for training to differentiate live skin textures from artificial ones. LBP is computationally efficient and robust to monotonic lighting changes, making it suitable for low-end cameras. It complements optical flow by providing a static analysis component. LBP may struggle with low-resolution images or extreme lighting conditions, where texture details are obscured. This is mitigated by combining it with optical flow and lighting analysis.

The system also analyzes natural lighting distribution to detect liveness by examining pixel intensity gradients across the face. Live faces reflect light naturally, creating smooth gradients influenced by facial contours, whereas static images often exhibit abrupt intensity changes due to reflections, edges of printed media, or screen glare. The system computes the gradient of pixel intensities using Sobel filters (horizontal and vertical) to capture the rate of change in brightness across the face. The gradient magnitude and direction are analyzed to detect smooth transitions typical of live faces versus sharp discontinuities in spoofed images. Live faces show consistent gradient patterns aligned with 3D facial geometry, while printed photos may display unnatural edges (e.g., paper borders) or screen images may show glare-induced spikes in gradient magnitude. A statistical measure, such as the variance of gradient magnitudes, is used to quantify these differences. The gradient analysis is performed on preprocessed 224x224 grayscale images, with a threshold on gradient variance to classify live versus spoofed faces. This is integrated into the CNN's liveness detection pipeline as an additional feature. Gradient analysis is robust to uniform lighting changes and complements optical flow by focusing on static image properties. Extreme lighting conditions (e.g., direct sunlight or low light) can distort gradients, necessitating adaptive thresholding.

The liveness detection features (optical flow, LBP histograms, and gradient statistics) are integrated into the CNN pipeline as follows: The CNN extracts facial features from preprocessed images using convolutional layers, while optical flow vectors, LBP histograms, and gradient statistics are computed separately and concatenated as additional input features to the fully connected layers. The CNN's final dense layer, with a softmax activation, outputs a binary classification (live vs. spoofed) based on the combined features. The model is trained end-to-end using a binary cross-entropy loss function, with the Adam optimizer and a learning rate of 0.001. The integration of these features contributes to the system's 91.3% accuracy, with optical flow being the dominant factor for micro-movement detection, as it captures dynamic behaviors most effectively. As noted, frame differences were not used due to their lack of directional information and sensitivity to noise. Optical flow provides a more robust representation of motion. While LBP is used for texture analysis, other methods like Histogram of Oriented Gradients (HOG) were considered but found to be less effective for fine-grained skin texture differentiation. Some liveness detection systems use depth sensors (e.g., Jee et al., 2006), but this requires specialized hardware. The proposed system's reliance on standard cameras makes it more accessible, though less precise for depth-based cues.

The liveness detection system is designed for standard webcams or mobile phone cameras, requiring at least 720p resolution and 10 fps for reliable optical flow computation. The system processes frames in real-time (approximately 50 ms per frame on the specified hardware: Intel Core i7-10700, NVIDIA RTX 3060). To handle varying conditions, the system applies adaptive thresholding and normalizes input data to mitigate noise from low-quality cameras or inconsistent lighting.

In summary, the liveness detection mechanism primarily relies on optical flow for micro-movement detection, complemented by Local Binary Patterns for texture analysis and pixel intensity gradient analysis for lighting evaluation. This multi-faceted approach ensures robust differentiation of live faces from spoofed images, addressing the core challenge of proxy attendance fraud in a software-based, hardware-agnostic manner.

3. Result and Discussions

This study focuses on creating a photo-based digital attendance system that detects fraudulent attempts by combining facial recognition with liveness detection capabilities. The system was tested using a dataset of 800 facial images, including valid inputs (captured in real-time) and invalid ones (such as static photos or images displayed on screens). Of the total dataset, 560 images were used for training, 120 for validation, and 120 for testing the model's effectiveness. The Convolutional Neural Network (CNN) model implemented is designed to distinguish between authentic and spoofed faces, with its performance validated through the following test outcomes:

Table 2. Results

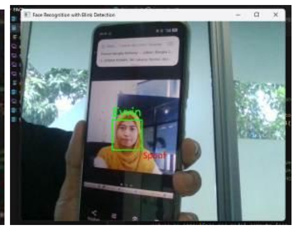
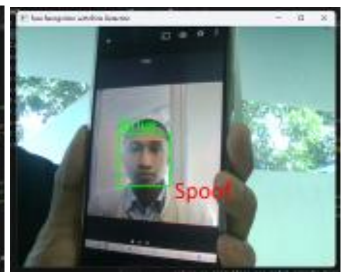
Subjects	Picture		
Sesa			
Evvin			
Sirlus			

Table 3. Confusion Matrix

Matrix	Value
Accuracy	91,3
Precision	89,6
Recall	90,1
F1-Score	89,8

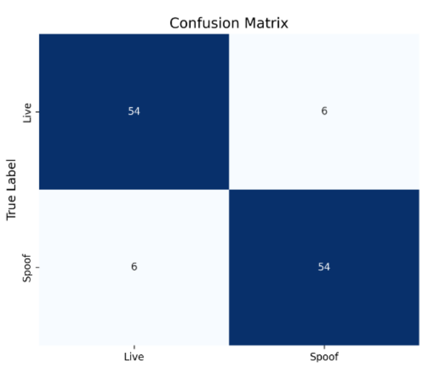


Figure 3. Confusion Matrix

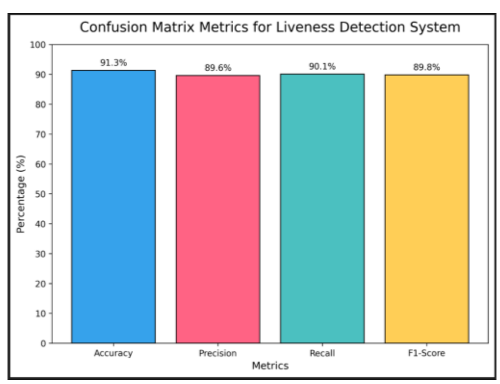


Figure 4. Confusion Matrix

The findings show that the system delivers robust classification performance, achieving a well-balanced precision and recall, indicating its effectiveness in distinguishing genuine from fraudulent faces. The main problem addressed in this study is the widespread issue of fraud in photo-based attendance systems, where individuals may use photos or screenshots of another person's face to falsely record attendance. To address this, a CNN-based model is combined with a liveness detection mechanism. Unlike conventional methods that depend on specialized hardware, this approach evaluates visual indicators such as skin texture, lighting distribution, and micro-movements (e.g., eye blinks or subtle head tilts) from static images or brief video clips.

The test results show that the model achieves high facial recognition accuracy while effectively differentiating between live faces and static images. This validates that integrating CNN with visual liveness detection provides a strong defense against basic spoofing attempts without requiring additional hardware. The practical outcome of this study is a secure, tamper-resistant attendance system ideal for academic and corporate settings where automated attendance tracking is essential without risking identity authenticity. Additionally, being entirely software-based, the system is easy to implement without necessitating significant changes to existing infrastructure. [21]

Deploying this system in real-world scenarios, such as educational institutions or corporate settings, introduces additional challenges that were not fully addressed in the controlled laboratory environment. One significant consideration is the system's performance under outdoor lighting conditions. Outdoor environments often involve dynamic lighting, such as direct sunlight, shadows, or overcast conditions, which can significantly affect the quality of captured images. The current system, which relies on analyzing pixel intensity gradients and skin texture, showed sensitivity to extreme lighting conditions (e.g., overly bright or dim environments). For instance, direct sunlight may cause overexposure, reducing the visibility of facial features, while low-light conditions can obscure micro-movements and texture details critical for liveness detection. To mitigate these issues, adaptive preprocessing techniques, such as histogram equalization or dynamic range adjustment, could be integrated to normalize images under varying lighting conditions.

Another critical factor is the reliance on low-end cameras, which are common in many real-world settings, particularly in resource-constrained environments like schools or small businesses. Low-end cameras, often with resolutions below 720p or limited frame rates, may fail to capture the fine details required for effective optical flow analysis or texture evaluation. The system's current requirement of at least 720p resolution and 10 fps may not be met by older mobile phones or budget webcams, potentially reducing accuracy. To address this, future iterations could incorporate super-resolution techniques or lightweight CNN architectures optimized for low-resolution inputs, as suggested by Horng et al. [22], who explored CNNs for small face images. Additionally, implementing frame interpolation to enhance temporal resolution could improve micro-movement detection on low-frame-rate devices.

The system's performance may also be affected by real-world factors such as diverse facial accessories (e.g., glasses, hats, or masks) and extreme facial expressions, which were identified as limitations in the study. In outdoor or uncontrolled settings, users may wear sunglasses or face coverings, further complicating facial recognition and liveness detection. To enhance robustness, the system could integrate multi-modal features, such as combining facial data with behavioral cues (e.g., head movement patterns), as proposed by Romeo et al. [1] for video-based analysis.

Real-world deployment also necessitates considerations for scalability and user accessibility. In large institutions with hundreds of users, the system must handle high throughput while maintaining real-time performance (currently ~50 ms per frame on high-end hardware). On low-end devices, processing times may increase, potentially causing delays in attendance recording. Optimizing the CNN model for edge devices, as explored by Khan et al. [19] for resource-constrained devices like smart glasses, could address this challenge. Additionally, ensuring compatibility with a wide range of mobile devices and operating systems is crucial for accessibility, particularly in diverse settings where users may rely on varying hardware.

Despite these challenges, the software-based nature of the system remains a significant advantage, as it eliminates the need for costly hardware upgrades. However, real-world testing in diverse environments—such as outdoor campuses or offices with mixed lighting and device types—is essential to validate the system's performance. Future research should focus on collecting larger, more diverse datasets that include outdoor images and low-quality camera inputs to improve model generalization.

The study identified several limitations, including the system's sensitivity to extreme lighting conditions (either too dark or excessively bright) and the quality of the camera used. Extreme facial expressions and accessories like masks also reduced the system's accuracy. Furthermore, testing was confined to controlled laboratory settings, leaving its performance in real-world conditions, such as outdoor environments or with low-quality cameras, unassessed. To overcome these challenges, future research should focus on developing a more robust system that leverages multi-frame or real-time video analysis. Incorporating motion-based or dynamic expression-based anti-spoofing methods and adaptive preprocessing for outdoor lighting is advised to improve resilience in varied and unpredictable settings. Additionally, optimizing the model for low-end devices and conducting real-world testing with diverse lighting and camera quality will enhance its practical effectiveness.

4. Conclusions and Future Works

This study successfully developed a photo-based digital attendance system that integrates CNN-based facial recognition with liveness detection, achieving a 91.3% accuracy rate in preventing fraudulent attendance [14]. The system mitigates security weaknesses in conventional photo-based systems by analyzing micro-movements, skin texture, and lighting patterns [3]. Its software-based design makes it well-suited for educational and public institutions [8][14]. However, challenges remain in extreme lighting conditions, with low-quality cameras, and with diverse facial expressions [4][22]. Future improvements should explore video-based methods, multi-modal authentication, and real-world testing with varied datasets to enhance system robustness [1][14].

The system effectively counters security flaws in traditional photo-based attendance systems, which are prone to exploitation. In addition to precise identity verification, it assesses visual authenticity through micro-level facial features, such as skin texture and natural lighting patterns. Nonetheless, it faces limitations in extreme lighting, low-resolution cameras, and highly variable facial expressions.

For future enhancements, adopting video- or multi-frame-based approaches is recommended to capture dynamic facial cues like blinking, head movements, or subtle expressions. Integrating additional technologies, such as multi-modal authentication combining facial recognition with voice or movement, could improve detection of advanced spoofing attempts. Furthermore, conducting real-world evaluations with diverse datasets is essential to boost the system's robustness and adaptability.

5. References

[1] L. Romeo, R. Marani, T. D'Orazio, and G. Cicirelli, "Video Based Mobility Monitoring of Elderly People Using Deep Learning Models," *IEEE Access*, vol. 11, pp. 2804–2819, 2023, doi: 10.1109/ACCESS.2023.3234421.

[2] D. Menotti *et al.*, "Deep Representations for Iris, Face, and Fingerprint Spoofing Detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 864–879, Apr. 2015, doi: 10.1109/TIFS.2015.2398817.

[3] Y. Wei, I. K. D. Machica, C. E. Dumdumaya, J. C. T. Arroyo, and A. P. Delima, "Liveness Detection Based on Improved Convolutional Neural Network for Face Recognition Security," *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, no. 8, pp. 45–53, Aug. 2022, doi: 10.46338/ijetae0822_06.

[4] A. Elmahmudi and H. Ugail, "Deep face recognition using imperfect facial data," *Future Generation Computer Systems*, vol. 99, pp. 213–225, Oct. 2019, doi: 10.1016/j.future.2019.04.025.

[5] A. Celecia, K. Figueiredo, M. Vellasco, and R. González, "A Portable Fuzzy Driver Drowsiness Estimation System," *Sensors*, vol. 20, no. 15, p. 4093, Jul. 2020, doi: 10.3390/s20154093.

[6] R. Mahum *et al.*, "A robust framework to generate surveillance video summaries using combination of zernike moments and r-transform and deep neural network," *Multimed Tools Appl*, vol. 82, no. 9, pp. 13811–13835, Apr. 2023, doi: 10.1007/s11042-022-13773-4.

[7] S. S. B. Benslet, P. Parameswari, and S. S. B. Benslet, “Criminal Facial Recognition Based on Multi Stage Progressive V-Net and MTCNN with NASnet Architecture,” 2025. [Online]. Available: https://www.jneonatalsurg.com

[8] K. Maisha and S. N. Shetu, “Influencing factors of e-learning adoption amongst students in a developing country: the post-pandemic scenario in Bangladesh,” *Future Business Journal*, vol. 9, no. 1, p. 37, Aug. 2023, doi: 10.1186/s43093-023-00214-3.

[9] A. Pandit, R. Nikalje, N. Vishwakarma, V. Vishwasrao, and Prof. T. Khose, “Face Authentication using MTCNN and FaceNet,” *Int J Res Appl Sci Eng Technol*, vol. 11, no. 11, pp. 1140–1143, Nov. 2023, doi: 10.22214/ijraset.2023.56679.

[10] J. A. Starosta and B. Izydorczyk, “Understanding the Phenomenon of Binge-Watching—A Systematic Review,” *Int J Environ Res Public Health*, vol. 17, no. 12, p. 4469, Jun. 2020, doi: 10.3390/ijerph17124469.

[11] C. Qin, X. Lu, P. Zhang, H. Xie, and W. Zeng, “Identity Recognition Based on Face Image,” *J Phys Conf Ser*, vol. 1302, no. 3, p. 032049, Aug. 2019, doi: 10.1088/1742-6596/1302/3/032049.

[12] F. N. Azmi, T. A. Z, and M. I. Suriansyah, “Sistem Keamanan Penerimaan Paket Menggunakan Face Recognition dan Metode Comparison Algorithms dengan Rasberry Pi Berbasis Internet of Things,” *J-INTECH*, vol. 12, no. 1, pp. 1–13, Jun. 2024, doi: 10.32664/j-intech.v12i1.1162.

[13] Y. Yanuardi, L. Azhari, A. A. J. Sinlae, and A. D. Alexander, “Pengembangan Sistem Pengaduan Layanan Masyarakat Menggunakan Metode Rapid Application Development (RAD),” *J-INTECH*, vol. 12, no. 1, pp. 36–48, Jun. 2024, doi: 10.32664/j-intech.v12i1.1201.

[14] R. E. Saragih and Q. H. To, “A Survey of Face Recognition Based on Convolutional Neural Network,” 2022.

[15] Z. Sun and Z. Liu, “Ensuring privacy in face recognition: a survey on data generation, inference and storage,” *Discover Applied Sciences*, vol. 7, no. 5, p. 441, May 2025, doi: 10.1007/s42452-025-06987-2.

[16] R. Jannah, M. Walid, and H. Hoiriyah, “Sistem Pengenalan Citra Dokumen Tanda Tangan Menggunakan Metode CNN (Convolutional Neural Network),” *Energy - Jurnal Ilmiah Ilmu-Ilmu Teknik*, vol. 12, no. 2, pp. 1–8, Dec. 2022, doi: 10.51747/energy.v12i2.1116.

[17] F. F. Wati, A. E. Widodo, and D. Abror, “Sistem Informasi Penerimaan Karyawan Berbasis Website Pada PT Federal International Finance Cabang Kota Tegal,” *J-INTECH*, vol. 12, no. 1, pp. 62–72, Jun. 2024, doi: 10.32664/j-intech.v12i1.1189.

[18] I. Ramadhan, N. Nugroho, H. Kurniawanto, and J. Warta, “Sistem Pendukung Keputusan Menggunakan Metode WASPAS Untuk Pemilihan Aplikasi Manajemen Bisnis dan Keuangan,” *J-INTECH*, vol. 12, no. 1, pp. 49–61, Jun. 2024, doi: 10.32664/j-intech.v12i1.1214.

[19] S. Khan, M. H. Javed, E. Ahmed, S. A. A. Shah, and S. U. Ali, “Facial Recognition using Convolutional Neural Networks and Implementation on Smart Glasses,” in *2019 International Conference on Information Science and Communication Technology (ICISCT)*, IEEE, Mar. 2019, pp. 1–6. doi: 10.1109/CISCT.2019.8777442.

[20] V. Musanga, C. Chibaya, and S. Viriri, “A scoping review of literature on deep learning and symbolic AI-based framework for detecting Covid-19 using computerized tomography scans,” *International Journal of Research in Business and Social Science (2147- 4478)*, vol. 13, no. 2, pp. 412–419, Apr. 2024, doi: 10.20525/ijrbs.v13i2.2955.

[21] A. Ahmed, J. Guo, F. Ali, F. Deeba, and A. Ahmed, “LBPH based improved face recognition at low resolution,” in *2018 International Conference on Artificial Intelligence and Big Data, ICAIBD 2018*,

Institute of Electrical and Electronics Engineers Inc., Jun. 2018, pp. 144–147. doi: 10.1109/ICAIBD.2018.8396183.

[22] P. Archana, N. Reddy, M. Bhavya, Y. Sharan, and M. Shashank, “Recognizing Very Small Face Images Using Convolution Neural Networks,” 2023.