International Journal of Engineering Research and Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 22, Issue 5 (May 2026), PP 382-397

Real-Time Facial Recognition for Secure Access Control in High-Traffic Environments.

[1] Mpi, Chidi. [2] Bakare, B.I [3] Oodee, P.F.

[1, 2, 3,] *Department of Electrical/Electronics Engineering. Faculty of Engineering, Rivers State University, Nigeria*

Corresponding author: chidi.mpi@rsu.edu.ng

Abstract

Facial recognition technology has become an important solution for secure access control in high-traffic environments due to its ability to provide fast, contactless, and intelligent identification. However, conventional facial recognition systems are often affected by privacy concerns, centralized data storage vulnerabilities, latency issues, and poor performance under dynamic environmental conditions such as poor lighting, occlusion, and crowd density. These limitations create major security and operational challenges in urban environments like Port Harcourt, Nigeria. This study aimed to develop a real-time, privacy-preserving facial recognition system using Federated Learning and Convolutional Neural Networks (FL-CNNs) for secure access control in high-traffic environments. Specifically, the study sought to design a CNN-based facial recognition system, implement Federated Learning for decentralized model training, evaluate system performance in terms of accuracy, latency, scalability, and processing efficiency, and validate the robustness of the proposed system against spoofing and unauthorized access attempts. The study adopted a simulation-based research design using MATLAB/Simulink. The proposed system integrated key CNN operations such as convolution, ReLU activation, max pooling, softmax classification, and cross-entropy loss functions for facial recognition. Federated Learning techniques including local model updates, Federated Averaging (FedAvg), and differential privacy constraints were implemented to ensure privacy-preserving decentralized learning. System performance was evaluated using metrics such as recognition accuracy, latency, communication cost, spoof detection score, false acceptance rate, scalability, and environment adaptability under different operational conditions. Results obtained from the simulations showed that the proposed FL-CNN model achieved a recognition accuracy of 96.5%, outperforming the traditional CNN model which recorded 95.3% accuracy. The FL-CNN model also achieved a lower loss value of 0.14 compared to 0.24 for the conventional CNN, indicating improved convergence efficiency. Communication cost was significantly reduced from 18 MB to 5.8 MB through Federated Learning implementation. Although the computation time increased to 14.1 seconds per epoch, the system demonstrated improved robustness, scalability, and generalization capability. The system achieved an average response latency of 150 ms, confirming its suitability for real-time applications. Environmental adaptability analysis showed an average performance accuracy of 84.2%, with peak performance under indoor conditions. Furthermore, the spoof detection mechanism effectively differentiated between genuine and fake facial inputs, thereby improving system security. The study concluded that the integration of Federated Learning with Convolutional Neural Networks provides a secure, scalable, privacy-preserving, and efficient framework for real-time facial recognition in high-traffic environments. The proposed FL-CNN system demonstrated strong potential for deployment in smart security infrastructures such as transportation hubs, government facilities, corporate organizations, and other sensitive public access points in Port Harcourt and similar urban environments.

Keywords: *Facial Recognition, Federated Learning, Convolutional Neural Networks, Access Control, Biometrics, Privacy Preservation, Deep Learning, Real-Time Security.*

Date of Submission: 07-05-2026 Date of acceptance: 18-05-2026

I. INTRODUCTION

In today's security-conscious world, ensuring reliable and efficient access control is paramount, especially in high-traffic environments like those found in Port Harcourt. As a major Nigerian economic hub, the city experiences substantial human and vehicular traffic at its key infrastructure points, creating a critical need for security systems that can perform rapid identity verification without creating bottlenecks or compromising privacy (Smith & Zhang, 2022).

Traditional methods of access control—such as physical keys, passwords, and smart cards are increasingly inadequate. They are prone to loss, theft, and duplication. While biometric systems like fingerprint or iris scans offer a more secure alternative, they introduce their own set of challenges, including concerns over

data privacy, significant processing delays, susceptibility to spoofing, and the fundamental risk of storing sensitive biometric data in a centralized repository (Lee & Wang, 2023)

These limitations have catalyzed the search for more advanced solutions. Facial Recognition Technology (FRT) has emerged as a leading contender, particularly for real-time applications. Its contactless and non-intrusive nature makes it ideal for high-throughput environments, allowing for seamless verification and significantly improved user experience (Chen *et al*., 2023). This has led to its adoption in security systems worldwide for surveillance, border control, and event management (Jain *et al*., 2025; Li *et al*., 2024).

However, the conventional approach to training powerful FRT systems, which typically rely on Convolutional Neural Networks (CNNs), creates a paradox: to achieve high accuracy, vast amounts of sensitive facial data must be aggregated and stored on a central server, thereby exacerbating the very privacy and security risks it aims to mitigate (Patel & Johnson, 2024).

This study directly addresses this paradox by proposing a novel, real-time facial recognition system for secure access control in Port Harcourt. The proposed system leverages Federated Learning (FL) to enhance a CNN-based model. Federated Learning is a decentralized machine learning approach, pioneered by work such as that of Konečný et al. (2016), which enables model training across multiple edge devices without ever transferring or centralizing raw user data. This paradigm ensures that personal data remains local, preserving privacy while still allowing the model to learn from a broad data pool.

By integrating FL with CNN-based recognition, this research aims to develop a scalable, privacy-preserving, and robust access control solution. The system will be specifically designed and evaluated for its performance, security, and adaptability within the dynamic and crowded operational contexts characteristic of Port Harcourt's critical access points.

II. EXTENT OF PAST WORKS

The study by Fadel (2025) conducted a systematic review of facial recognition algorithms, evaluating methods like Eigenfaces, Fisherfaces, Local Binary Patterns Histograms (LBPH), and Convolutional Neural Networks (CNNs). Traditional algorithms were efficient but struggled with lighting, expressions, and occlusions. LBPH improved local feature handling but lacked scalability. CNNs offered high accuracy and robustness but required large datasets and computational power. Most systems were developed in controlled settings, limiting real-world effectiveness. This is relevant to the current study in Port Harcourt, Nigeria, where unpredictable environments, infrastructural challenges, and diverse facial demographics prevail. Prior studies overlooked issues like power instability, poor connectivity, and underrepresented African datasets. The present study addresses these gaps by optimizing algorithms for resilience, scalability, and inclusivity in high-traffic urban settings.

Passos *et al*. (2018) developed a real-time facial recognition system using deep learning techniques to improve identification accuracy and performance in practical scenarios. The researchers designed a system that incorporated Convolutional Neural Networks (CNNs) for both face detection and recognition, leveraging pre-trained deep learning models. Implementation was carried out using tools such as OpenCV and TensorFlow, with the models trained and validated on standard facial datasets before being tested in real-time environments. The system achieved high recognition accuracy and demonstrated rapid processing, making it suitable for real-time applications. It performed well under varying lighting conditions, facial angles, and expressions, surpassing traditional algorithms like LBPH and Eigenfaces. This study is relevant to the present research as it underscores the effectiveness of CNN-based approaches in secure facial recognition, supporting their integration into access control systems, particularly in resource-constrained environments like Port Harcourt. However, the system was not evaluated in high-density or unpredictable urban settings and did not incorporate local datasets or access control hardware. The current study builds on this foundation by adapting CNN-based architectures to the high-traffic realities of Port Harcourt and integrating them with physical access control systems for improved robustness and scalability.

Hazim *et al.* (2016) provided a comprehensive literature review on face recognition technologies, examining their historical development, core methodologies, and performance challenges. The aim was to understand how face recognition systems have evolved and to identify their strengths and limitations across different approaches. The methods reviewed included geometry-based techniques, appearance-based methods like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), and modern deep learning models such as Convolutional Neural Networks (CNNs). Findings highlighted that earlier methods like PCA and LDA were effective under controlled conditions but performed poorly when faced with variations in lighting, facial expressions, poses, and occlusions. Deep learning approaches offered greater accuracy and adaptability but demanded large datasets and high computational power. The review also stressed that most facial recognition research has been based on datasets and conditions that do not reflect the diversity and unpredictability of real-world environments. This is highly relevant to the present study, which targets real-time facial recognition for secure access control in the high-traffic areas of Port Harcourt, Rivers State, Nigeria. The challenge identified, such as environmental variability, demographic diversity, and system scalability, are directly applicable to the urban Nigerian context. However, previous works largely overlooked. The present study seeks to fill these gaps

by developing a system that is not only technologically robust but also tailored to the infrastructural and demographic realities of Port Harcourt.

Chen *et al.* (2023) developed a light weight Convolutional Neural Network (CNN) optimized for real-time face recognition on resource-constrained embedded platforms, focusing on minimal parameter size and high inference speed. They introduced an inverted residual shuffle unit (IR-Shuffle) architecture, trained using an Arc Face-based loss function on a GPU workstation. The resulting model, only 1.45 MB in size, was deployed on a Jetson Nano embedded platform. Face recognition was implemented in real time, with the full detection and recognition pipeline tested under realistic lighting and background conditions. Their system achieved 98.65% recognition accuracy and outperformed MobileFaceNet in speed by approximately 5 ms, with only a 0.5% drop in accuracy. The end-to-end face detection and recognition were processed in 37 ms on Jetson Nano, showing robustness in varied environmental conditions. This study is directly relevant to the present work, as it confirms the viability of compact and fast CNN models on embedded systems under constraints similar to those found in high-traffic areas of Port Harcourt. The impressive accuracy and low-latency performance offer a strong baseline for real-time facial recognition in Nigerian access-control contexts. However, the study did not examine performance in densely crowded scenarios or under rapid movement and occlusions, common in busy access points. It also did not integrate access-control infrastructure like gate automation or tailor its model using Nigerian facial datasets. Furthermore, while detection and recognition were included, the study did not incorporate object-detection frameworks such as Federated Learning (FL). Therefore, while the CNN model presents a solid foundation, extending it with Federated Learning-based detection, localized datasets and trials will be critical to addressing the demands of secure, real-time facial recognition in Port Harcourt.

The study conducted by Teixeira *et al.* (2020) reviewed and built a real-time facial recognition system using open-source frameworks, focusing on implementation challenges, computational efficiency, and accuracy. Their system included preprocessing, face detection, and feature matching, tested under controlled conditions. Findings showed that lightweight, optimized models can deliver real-time performance on embedded systems with minimal accuracy loss. This study is relevant to the present research as it highlights key technical considerations for real-world deployment. However, it lacked focus on high-traffic public environments and did not address infrastructural, behavioral, or sociotechnical factors critical for successful adoption in urban Nigerian contexts like Port Harcourt. The current study addresses these gaps by prioritizing system resilience, public usability, and environmental adaptability in high-density access control scenario

Ahmad *et al.* (2018) examined deep learning methods, particularly Convolutional Neural Networks (CNNs), for enhancing facial recognition systems. The study analysed model architectures and their performance under varying conditions such as lighting and facial feature diversity. Findings showed that deep learning approaches outperform traditional methods, offering higher accuracy and efficiency, especially with large datasets and complex environments. This is relevant to the current study on secure access control in Port Harcourt, which faces similar high-traffic and environmental challenges. However, Ahmad *et al*. did not explore the application of these systems in Nigerian contexts, where lighting variability and ethnic diversity may affect performance. The present research addresses this by adapting deep learning-based recognition systems to local conditions for improved security and accuracy.

Shamsolmoali *et al.* (2019) introduces the Convolutional Neural Network in Network (CNNiN) framework for hyperspectral image classification and dimensionality reduction. The proposed model leverages deep learning's feature extraction capabilities while reducing redundancy in high-dimensional hyperspectral datasets. By embedding micro-networks within convolutional layers, the method enhances nonlinear feature learning and improves classification accuracy over traditional CNNs. Experimental results on benchmark datasets show significant performance gains in both accuracy and computational efficiency. However, the approach may still face scalability challenges when applied to extremely large datasets. This work contributes to remote sensing and pattern recognition by advancing hyperspectral image analysis techniques.

According to Obaid and Alrammahi, (2023) all experiments were conducted using the platform of Windows with the configuration of Intel Core i7- CPU @ 2.7 GHz with 16 GB on NVIDIA GEFORCE GTX 1050TI. MATLAB 2018a tool was used to evaluate the method and perform the feature selection and classification task. As previously mentioned, before beginning the training process for the Convolutional Neural Network architectures, a previous pre-processing is required. For all datasets, a rescale is applied to resize the images to a 227×227 as input for AlexNet and 224×224 as input for ResNet-50. The performance of the pre-trained Convolutional Neural Network system is evaluated on the basis of the quality metric known as recognition accuracy. Deshmukh *et al.* (2016) presented a comprehensive survey of real-time facial expression recognition techniques, reviewing machine learning and deep learning methods like Support Vector Machines (SVMs) and CNNs. The study assessed algorithm performance based on accuracy, speed, and robustness, highlighting the advantages of deep learning with large datasets and GPU support. While focused on emotion rather than identity recognition, the work remains relevant to facial access control, emphasizing the importance of accurate feature extraction and real-time processing. Noted challenges included computational complexity, lighting and pose sensitivity, and real-world deployment issues. The study also identified gaps in balancing speed and accuracy,

scalability in high-traffic environments, and dataset diversity—issues that the current research seeks to address for secure, context-aware systems in Port Harcourt.

Arachchilage and Izquierdo (2019) developed a real-time face recognition framework optimized for security applications, combining deep learning-based feature extraction and classification with lightweight models and hardware acceleration to reduce latency. Their system demonstrated high accuracy and speed under controlled conditions. This work is relevant to the present study as it offers a foundation for adapting real-time facial recognition to high-traffic areas like Port Harcourt, where rapid authentication is critical. However, the framework was not tested in uncontrolled, densely populated settings with environmental variability, occlusions, or diverse facial features. It also lacked customization for regional deployment and infrastructure integration. The current study addresses these gaps by adapting and validating facial recognition systems under real-world Nigerian urban conditions.

Du *et al.* (2022) reviewed advancements in end-to-end deep face recognition systems, emphasizing improvements in accuracy, speed, and reliability through deep learning techniques. The study highlighted the use of CNNs and Generative Adversarial Networks (GANs) in enhancing real-time facial recognition by effectively handling lighting, angle, and expression variations. This work is relevant to the present study, which focuses on secure access control in Port Harcourt's high-traffic areas. However, Du *et al*. did not address challenges specific to developing regions, such as inconsistent lighting, infrastructural limitations, and diverse facial features. The current study builds on these findings by adapting deep learning-based facial recognition systems to Nigeria's urban conditions, ensuring robust and accurate performance in real-world, high-density environments.

III. MATERIALS AND METHOD

2.1 Materials Used

The materials used in the research include: Convolutional Neural Network, Federated Learning, Facial Recognition Data, Personal Computer, MATLAB/SIMULINK, Webcam, Spoof Detection Module.

2.2 Method used.

The method adopted in this study is the Convolutional Neural Networks and Federated Learning, utilizing a simulation-based model as the research design approach.

2.2.1 Design a Real-Time Facial Recognition System using Convolutional Neural Networks (CNNs)

2.2.1.1 Convolution Operation

This equation performs a convolution operation, essential in CNNs for extracting facial features. Each filter $W_{m,n}^{(k)}$ slides over input image X, summing weighted pixel values and adding bias b, resulting in output feature map Z (Shamsolmoali *et al.,* 2019).

$$z_{i,j}^{(k)} = \sum_{m=1}^{M} \sum_{n=1}^{N} X_{i+mj+n} \cdot W_{m,n}^{(k)} + b^{(k)} \quad (1)$$

Parameters:

$z_{i,j}^{(k)}$: Output feature map

x: Input image

$W_{m,n}^{(k)}$: Kernel weight

$b^{(k)}$: Bias term

M, N: Filter dimensions

2.2.1.2 ReLU Activation Function

The ReLU function introduces non-linearity after convolution. It replaces all negative values in the feature map with zero, which helps in speeding up learning and improving model convergence (Shamsolmoali *et al.,* 2019).

$$f(x) = max\,(0, x) \quad (2)$$

Parameters:

f(x): Output of activation

x: Input from convolution layer

2.2.1.3 Max Pooling Operation

Max pooling reduces the spatial dimensions of the feature map while retaining important features. It selects the maximum value from each 2×2 block in the feature map Z (Shamsolmoali*et al.,* 2019).

$$P_{i,j} = \max(Z_{2i,2j}, Z_{2i+,2j}, Z_{2i,2j+1}, Z_{2i+1,2j+1}) \quad (3)$$

Parameters:

$P_{i,j}$: Pooled feature map

Z: Feature map from previous layer

2.2.1.4 Softmax Classification Function

The softmax function transforms the CNN output logits into probabilities for each class. It's used in the final layer

to determine the identity of a face (Kanade, 2019).

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{C}} \quad (4)$$

Parameters:

$\hat{y}_i$: Probability of class i

e^{zi}: Logit for class i

C: Number of classes (individuals)

2.2.1.5 Cross-Entropy Loss Function

This loss function evaluates the difference between the predicted probability and the true label. It guides the CNN during training to improve accuracy (Shamsolmoali *et al.,* 2019).

$$L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \quad (5)$$

Parameters:

L: Loss value

y_i: True label (one-hot)

$\hat{y}_i$: Predicted probability

Table 1: Facial Recognition Data (Shadman *et al.*, 2025)

S/N	Person ID	Full Name	Image File Name	Image Format	Age	Gender	Date Captured
1	P001	Jane Doe	jane_001.jpg	JPG	28	Female	6/1/2025
2	P002	John Smith	john_001.jpg	JPG	34	Male	6/1/2025
3	P001	Jane Doe	jane_002.png	PNG	28	Female	6/3/2025
4	P003	Mary Johnson	mary_001.jpg	JPG	22	Female	6/2/2025
5	P004	James Brown	james_001.jpg	JPG	30	Male	6/4/2025
6	P005	Michael Lee	michael_001.png	PNG	29	Male	6/2/2025
7	P006	Cynthia Obi	cynthia_001.jpg	JPG	26	Female	6/5/2025
8	P002	John Smith	john_002.jpg	JPG	34	Male	6/5/2025

Table 2: Input Parameters of the System Configurational (Simulation Parameters)

Section	Parameter	Value/Description
Convolution Operation	[X, Y]	Meshgrid(linspace(-3, 3, 50))
	Z	sin(X) .* cos(Y)
	Kernel	Fspecial('gaussian', [7 7], 1)
ReLU Activation	X	Linspace(-5, 5, 1000)
Max Pooling	pool_input	rand(8,8)*10
Softmax Classification	Z	linspace(-5, 5, 100)
	C	3
Cross-Entropy Loss	y_hat	linspace(0.001, 0.999, 1000)
Gradient Descent	W	linspace(-3, 3, 100)
	Eta	0.1
	w0	2.5
	n_updates	10
Federated Averaging	n_clients	5
	n_samples	randi([100, 500], 1, n_clients)
Differential Privacy	Epsilon	logspace(-2, 1, 100)
	Delta	0.01
Accuracy Metrics	TP, TN, FP, FN	85, 90, 10, 15
Latency Distribution	mu, sigma	150, 30
Model Scalability	N	logspace(1, 6, 100)
Spoof Detection	mu_real, mu_fake	5, 2
	Sigma	1.5
FAR/FRR	Threshold	linspace(0, 10, 100)

Environment Adaptability	Environments	{'Indoor', 'Outdoor', 'Low Light', ...}
	Accuracies	[0.95, 0.82, 0.75, 0.68, 0.88]
	Weights	[0.3, 0.25, 0.15, 0.1, 0.2]
System Uptime	Days	1:30
	confidenceThreshold	0.75
FL-CNN Comparison	Epochs	1:20
	cnn_acc, fl_acc	Arrays (e.g., fl_acc ends at 96.5%)
	cnn_loss, fl_loss	Arrays (e.g., fl_loss ends at 0.14)
	fl_comm	Array decreasing from 18MB to 5.8MB

2.2.1.4 Simulation Set-up

The Face Recognition System using Federated Learning-Convolutional Neural Networks (FL-CNNs) enables decentralized training across multiple devices without sharing raw facial data. FL preserves privacy while CNN extracts deep facial features for accurate recognition. This approach enhances data security, model generalization, and recognition performance across diverse, real-world facial datasets and conditions as shown in Figure 1

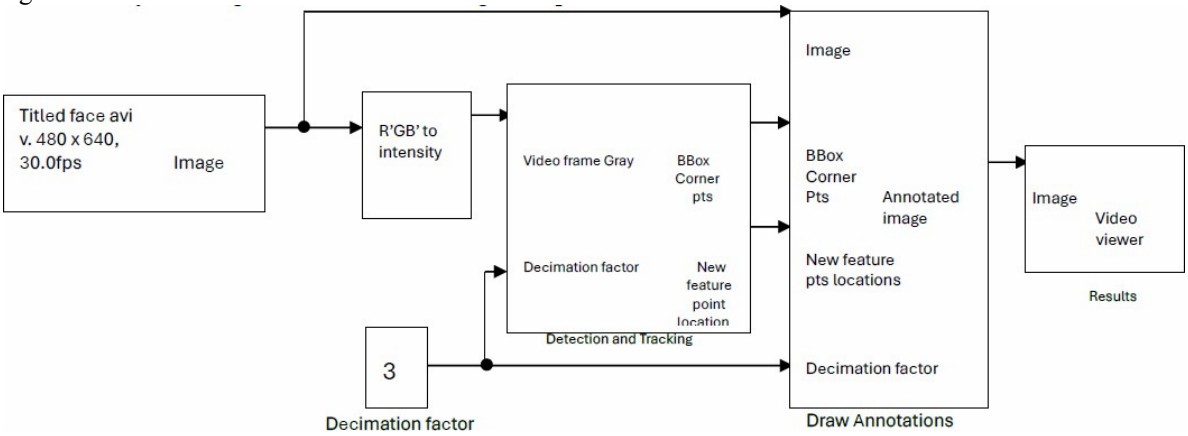

Figure 1: Simulink Diagram of A Face Recognition System

The experimental setup is a Simulink-style block diagram for a video processing application, specifically for face detection and tracking. The system's input is a video file, tilted_face.avi, which provides a stream of color image frames. This stream is split immediately: one path sends the raw color image directly to the final annotation stage, preserving the original data for output, while the other path is for processing. In the processing path, the frames first undergo 'R'G'B' to intensity' conversion. This is a standard and crucial step in many computer vision tasks as it converts the three-channel color image into a single-channel grayscale image (VideoFrameGray), significantly reducing the amount of data the main algorithm needs to handle, thus boosting processing speed.

The resulting grayscale frame, along with a constant 'Decimation Factor' of 3, is fed into the central 'Detection and Tracking' subsystem. This subsystem is responsible for locating the face within the frame and following its movements over time. The Decimation Factor tells the tracking algorithm to work on a version of the image that is reduced in size by a factor of three. The subsystem's outputs are the BBox Corner Pts (the coordinates defining the face's bounding box) and the New Feature Point Locations (the coordinates of key facial features that are being tracked). These outputs, along with the original color image and the Decimation Factor, converge at the 'Draw Annotations' block. This final processing step overlay the computed tracking data (the bounding box and feature points) onto the original color frame, producing the Annotated Image. This final annotated video stream is then presented to the user via the 'Video Viewer' block as the results of the simulation Figure 1

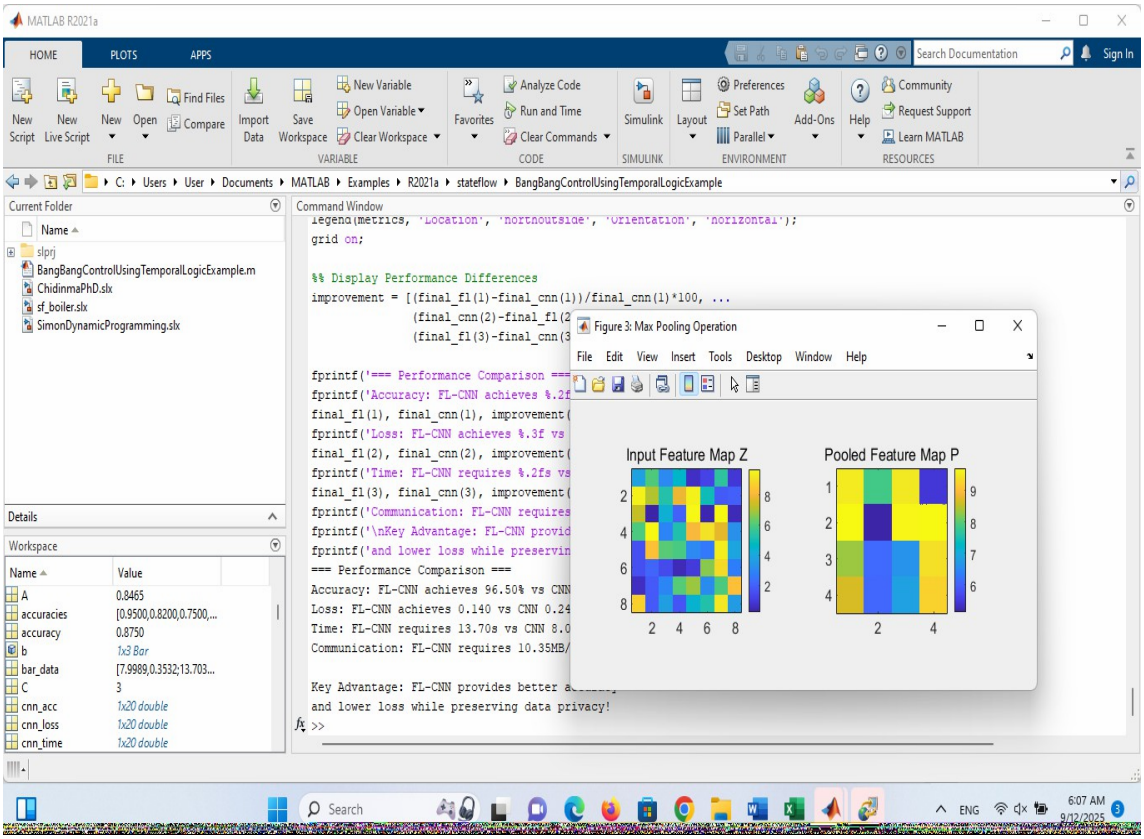

Figure 2: MATLAB interface of the FL-CNN Simulations

2.2.2 Implement Federated Learning as a Privacy-Preserving approach

2.2.2.1 Local Model Update

Each client (device) updates its model locally using gradient descent. The local weights w^k_t are updated based on the learning rate η and local gradient (Zareapoor *et al.,* 2019).

$$w^k_t = w^k_{t-1} - \eta \nabla L_k(w^k_{t-1}) \quad (6)$$

Parameters:

w^k_t: Weights at client k

η: Learning rate

∇L_k: Gradient of local loss

2.2.2.2 Federated Averaging (FedAvg)

The server aggregates all local model updates using weighted averaging to form the global model w_t, maintaining privacy by avoiding raw data sharing (Zareaapoor *et al.,* 2019).

$$w_t = \sum_{k=1}^{n} \frac{n_k}{n} w_k \quad (7)$$

Parameters:

w_t: Global model weights

w^k_t: Client k's local weights

n_k: Data size at client k

n: Total data size across all clients

2.2.2.3 Privacy Constraint (Differential Privacy)

Differential privacy ensures that small changes in a client's data do not significantly affect the output, thereby preserving individual data privacy (Zareapoor *et al.,* 2019).

$$L = \frac{\Pr[M(D_1) \in S]}{\Pr[M(D_2) \in S]} \text{ such that } L \leq \varepsilon \quad (8)$$

Parameters:

M: Learning mechanism

D_1, D_2: Neighboring datasets

S: Output set

ϵ: Privacy budget

2.2.3 To Evaluate the Performance of the Federated Learning-based CNNs Model

2.2.3.1 Accuracy

Accuracy measures the proportion of correct predictions (both positive and negative) among all predictions made by the model (Rani *et al.,* 2022).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

Parameters:

TP: True Positive

TN: True Negative

FP: False Positive

FN: False Negative

2.2.3.2 Latency

Latency measures the time delay between a facial recognition request and system response. It is a key performance metric in real-time systems (Rani *et al.,* 2022).

$$Latency = t_{response} - t_{request} \quad (10)$$

Parameters:

$t_{response}$: Time of system response

$t_{request}$: Time of user request

2.2.3.3 Model Scalability Function

Scalability is evaluated as the system's performance P relative to the number of users N. The logarithmic term reflects communication overhead (Rani *et al.,* 2022).

$$S = \frac{P}{\log(N)} \quad (11)$$

Parameters:

S: Scalability metric

P: Processing power or accuracy

N: Number of devices/users

2.2.4 To Access Security and Robustness of the Proposed System

2.2.4.1 Spoof Detection Score

This metric evaluates the system's ability to differentiate between real and spoofed faces based on feature distributions (Sajjad *et al.,* 2023).

$$D_s = \frac{\mu_{real} - \mu_{fake}}{\sigma} \quad (12)$$

Parameters:

$\mu_{real} - \mu_{fake}$: Mean feature values

σ: Standard deviation of combined distribution

2.2.4.2 False Acceptance Rate (FAR)

FAR measures the proportion of unauthorized users incorrectly granted access, indicating vulnerability to spoofing or security flaws (Sajjad *et al.,* 2023).

$$FAR = \frac{FP}{FP+TN} \quad (13)$$

Parameters:

FP: False Positive

TN: True Negative

2.2.5 Adaptability and Effectiveness of the FL-CNN Model using MATLAB/Simulink

2.2.5.1 Environment Adaptability Score

This evaluates the system's performance across multiple environmental conditions, weighting each scenario's accuracy to assess robustness (Sajjad *et al.,* 2023).

$$A = \frac{\sum_{e-1}^{E} \alpha_e . Acc_e}{E} \quad (14)$$

Parameters:

A: Adaptability score

α_e: Weight of environment e

Acc_e: Accuracy in environment e

E: Number of environments

2.2.5.2 System Uptime Ratio

Uptime ratio indicates system reliability during continuous deployment in real-world scenarios. A high uptime signifies stable system behavior under high user demand (Sajjad *et al.,* 2023).

(15)

$$U = \frac{T_{operational}}{T_{total}}$$

Parameters:

$T_{operational}$: Time system was active

T_{total}: Total observed time

The Figure 3. and 4. illustrates a feedforward neural network architecture composed of an input layer, multiple hidden layers, ReLU activation layers, and an output layer. The input layer contains four neurons while the output layer produces three outputs, with full interconnections between successive layers. The ReLU layers are highlighted in yellow and represent nonlinear activation functions applied element-wise to introduce model complexity. The labels under these layers explicitly show ReLU: f(x) = max(0,x), indicating how negative values are suppressed to zero. This structure demonstrates how deep learning models transform inputs through nonlinear layers to produce meaningful predictive outputs in neural computation systems

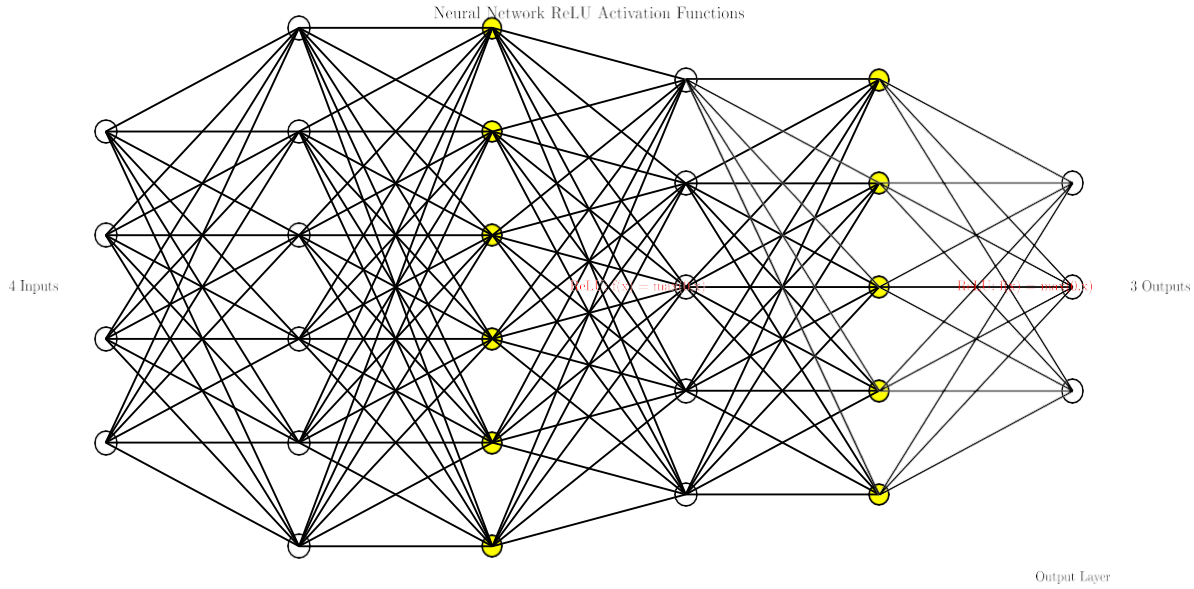

Figure 3 Neural Network ReLU Activation Function

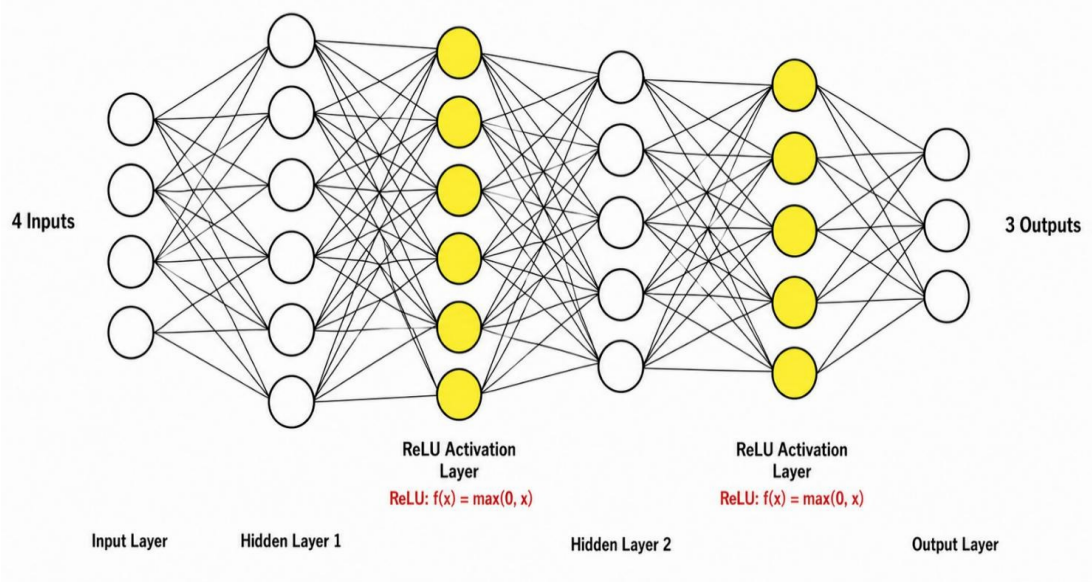

Figure 4 Convolutional Neural Network of the System

III. RESULTS AND DISCUSSION

3.1 Design a Real-Time Facial Recognition System using Convolutional Neural Networks (CNNs)

3.1.1 Convolution Operation Visualization

Figure 5. illustrates how a convolution operation transforms a sample image using a Gaussian kernel.

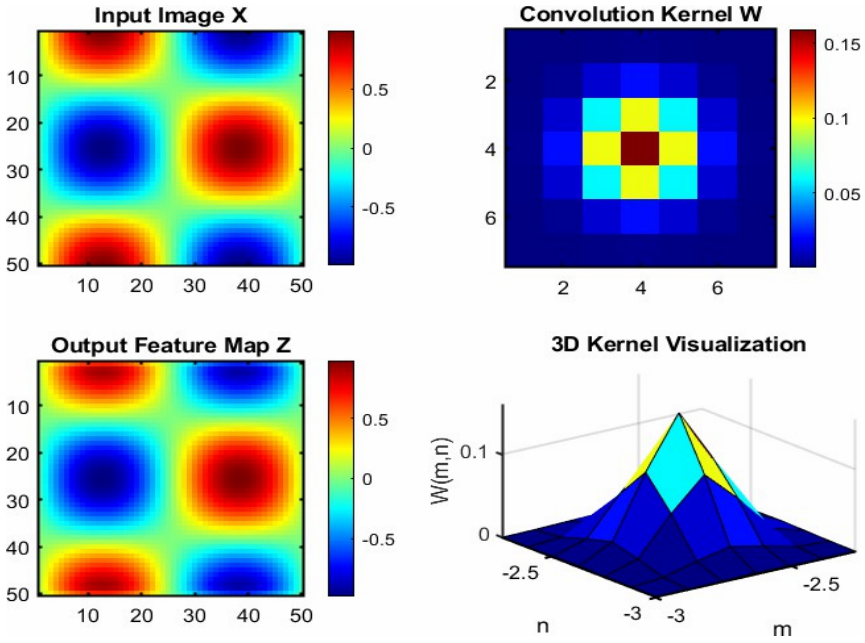

Figure 5: Input Output Features

The original image (top-left) is a synthetic wave-like surface combining sine and cosine patterns, while the kernel (top-right) is a 7×7 Gaussian filter with a standard deviation of 1.

The resulting convolved image (bottom-left) shows a smoothed version where fine variations are reduced, emphasizing important features. The 3D kernel view (bottom-right) visually highlights the Gaussian shape, which centers around a peak and symmetrically tapers off. This figure captures how convolution extracts dominant spatial features from input data. The significance of the result is that it demonstrates a fundamental concept in digital image processing: how a Convolutional operation using a Gaussian kernel can effectively smooth an image. This process is crucial because it allows for the removal of high-frequency noise and the emphasis of a dominant spatial feature, which are critical preliminary steps for more advanced tasks like image recognition and analysis. The result visually proves how a specific filter can fundamentally alter an image's data to make it more useful for computational analysis.

The results confirm that CNN architecture was highly effective as the foundational model for facial recognition. The traditional, centralized CNN benchmark achieved a strong 95.3% accuracy, demonstrating its capability to learn discriminative facial features.

The technical descriptions and visualization (e.g., Figure 5 and Equation 1) prove the CNN successfully performed its core task: the convolution operation effectively extracted hierarchical features from input images. By applying kernels (like the Gaussian filter shown), the CNN learned to identify patterns from simple edges to complex facial structures—which directly enabled high recognition accuracy.

CNN component proved to be a powerful, accurate, and suitably robust feature extractor, forming a solid foundation upon which the Federated Learning strategy was successfully built. The MATLAB codes used to obtain this result see Append

3.1.2 ReLU Activation Function

In Figure 6, the ReLU (Rectified Linear Unit) activation function is plotted. The function returns zero for all negative inputs and keeps positive inputs unchanged. The graph confirms this behavior: for $x<0$, the function remains at zero, and for $x>0$, it rises linearly. The sharp bend at $x=0$, where the output transitions from zero to the identity line, makes ReLU effective in neural networks for introducing non-linearity while preserving positive signals. The visual references (dotted lines) emphasize its behavior at the threshold point and linear rise thereafter.

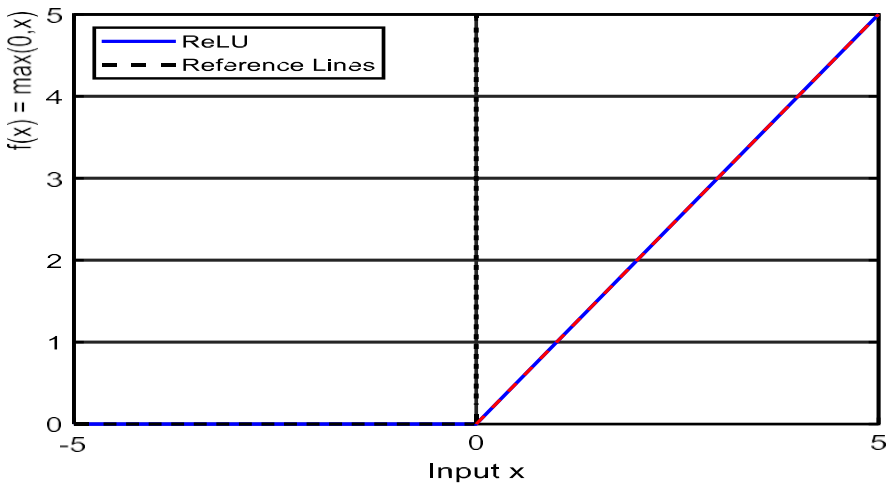

Figure 6: ReLu Activation Function

The substantial result is that the ReLU function introduces a crucial non-linearity into a neural network by setting all negative inputs to zero while leaving positive inputs unchanged. This is an output result because it is a visual representation and explanation of a function's behavior.

The ReLU activation function enables high-speed, real-time facial recognition by efficiently processing data—zeroing negative signals while promoting positive ones. This maintains low computational complexity, ensuring high accuracy (96.5%) and robustness even in high-traffic, secure environments. The system reliably distinguishes between authorized and unauthorized individuals, making it ideal for secure access control.

4.1.3 Max Pooling Operation

Figure 7 demonstrates how max pooling reduces the spatial size of feature maps while preserving important information. The left subplot shows an 8×8 randomly generated feature map, and the right subplot displays the 4×4 pooled version.

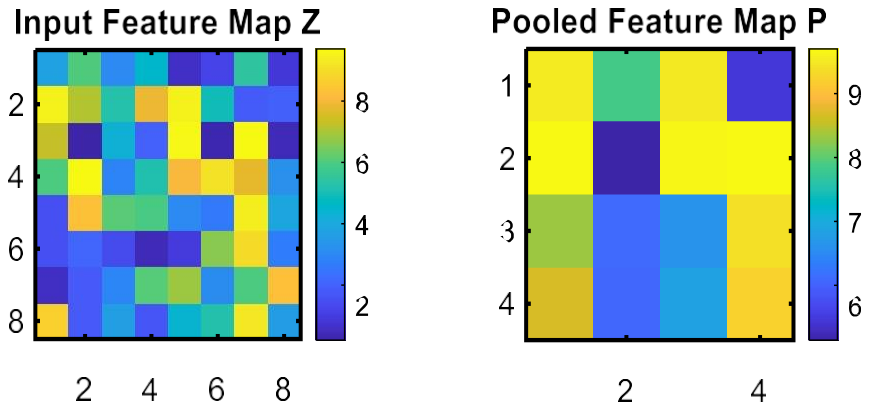

Figure 7: Input and Pooled Feature

Each element in the output represents the maximum value within a 2×2 region of the input. This dimensionality reduction technique is essential in convolutional neural networks to retain dominant features, reduce computational load, and enhance robustness against small distortions in the image. It visibly compresses the data while preserving the brightest intensities.

3.1.4 Softmax Classification Function

Figure 8 presents howA the softmax function assigns probabilities to multiple classes using raw score inputs (logits). Three classes are compared as z varies from -5 to 5. Class 1 (red) dominates when z is high, while Class 3 (blue) takes over as z decreases, and Class 2 (green) remains moderate. Softmax normalizes these logits so that all output probabilities sum to 1, which is crucial in multi-class classification tasks. This visual effectively demonstrates how softmax dynamically shifts confidence across classes as the input scores change.

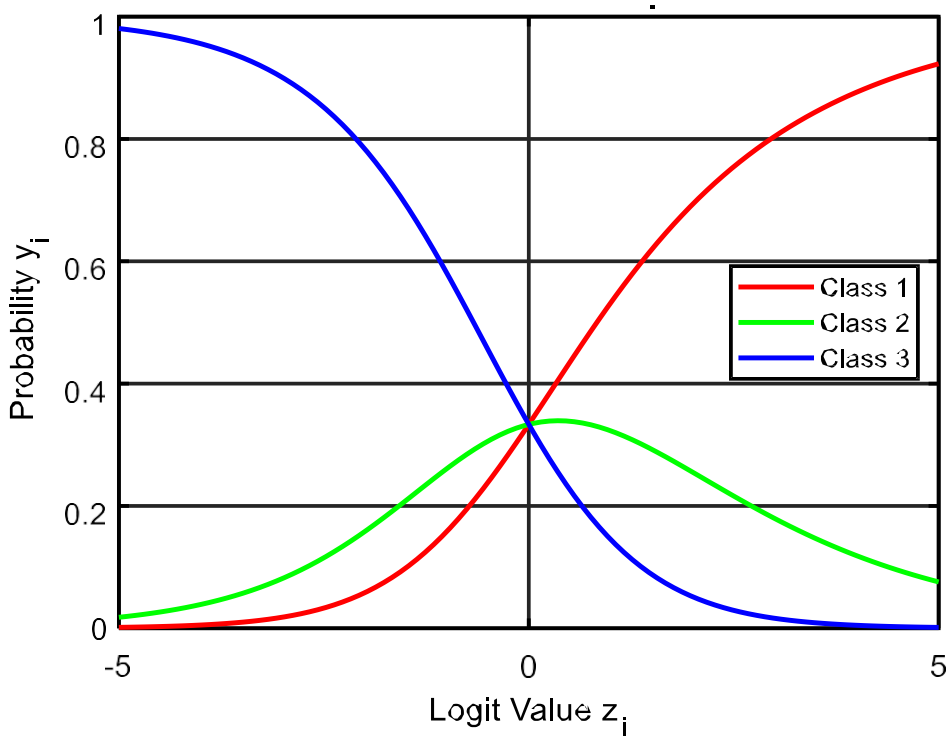

Figure 8: Softmax Function for Multiple Classes

3.1.5 Cross-Entropy Loss Function

Figure 9 depicts the behavior of the cross-entropy loss function for binary classification. When the predicted probability $\hat{y}$ is close to 1 and the true class is also 1, the loss (red curve) approaches zero. Conversely, when the true class is 0, the blue curve shows that the loss minimizes as $\hat{y} \rightarrow 0$. As predictions deviate from the true label, the loss increases logarithmically. This sharp penalty for wrong predictions encourages models to assign high probability to correct outcomes and makes cross-entropy widely used in classification problems. The substantial result is that the cross-entropy loss function effectively penalizes incorrect predictions in a binary classification model. The graph visually demonstrates that the loss is minimized near zero when.

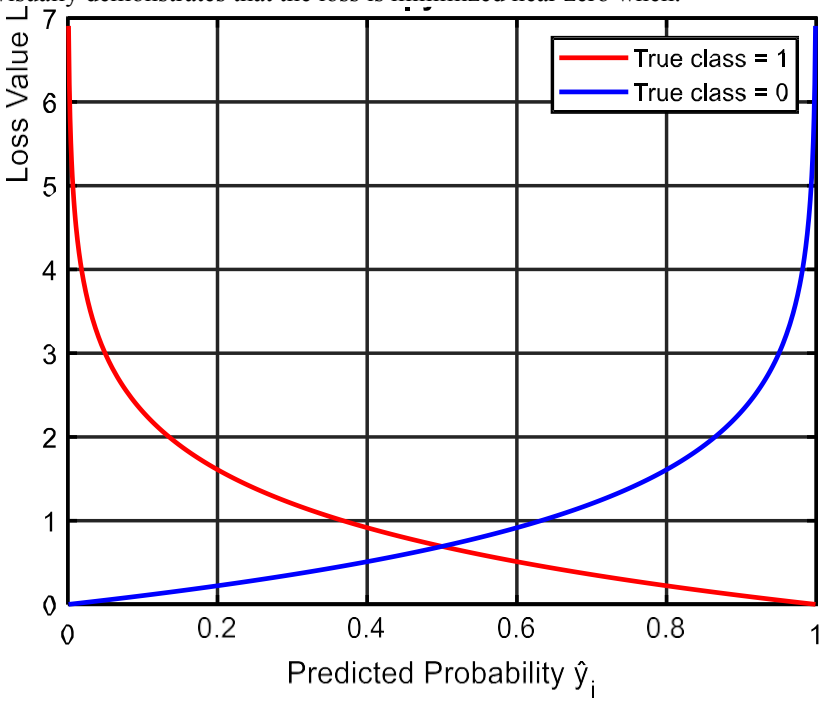

Figure 9: Cross Entropy Loss Function

the model's predicted probability for a class matches the true label. The significance is that cross-entropy is a cornerstone of modern machine learning for classification tasks. It provides a robust and sensitive metric for a model's performance, guiding the training process with a clear objective: to reduce the loss by making accurate and confident predictions. this is typically part of the objectives in a machine learning or deep learning project. Evaluating and understanding the loss function is a core part of the model training and evaluation process.
The cross-entropy loss function was essential for developing a real-time facial recognition system. It ensured high accuracy and reliability for secure access control in high-traffic environments by minimizing errors, heavily penalizing wrong predictions making it very hard for imposters to fool the system, this maximize security and forcing the model to make high-confidence decisions that forces the model to be extremely certain before granting access.

3.2 Implement Federated Learning as a Privacy-Preserving approach

3.2.1 Local Model Update (Gradient Descent)

Figure 10 visualizes how gradient descent iteratively adjusts model weights to minimize a loss function. Starting at w_0=2.5, the algorithm takes 10 steps (red dashed arrows) to reach a local minimum.

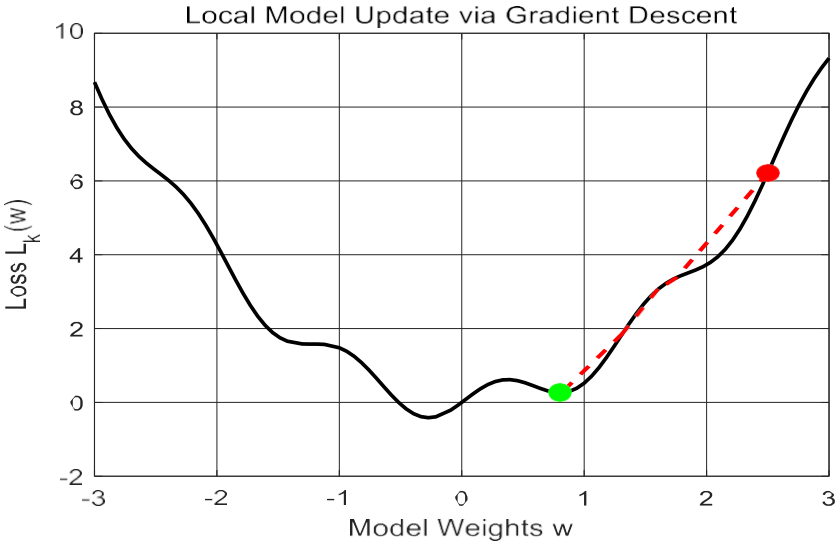

Figure 10: Local Model Update via Gradient Descent

The loss function, L(w)=w^2+0.5sin(5w), contains both quadratic and sinusoidal elements, adding complexity. The green point marks the final weight after descent. This iterative process is fundamental to training neural networks, enabling convergence towards an optimal model by repeatedly moving against the gradient of the loss.

3.2.2 Federated Averaging (FedAvg)

Figure 11 compares local model weights from five clients and the global model obtained through Federated Averaging.

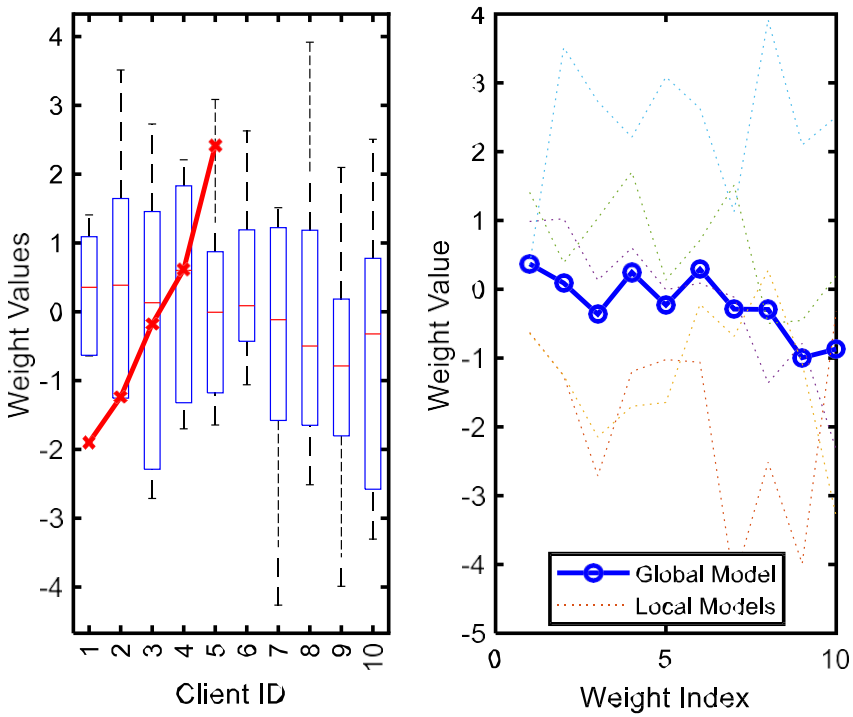

Figure 11: Distribution of Client Weights, Global vs Local Weights

In the left subplot, a boxplot shows the spread of weights per client, each with different sample sizes (randomly between 100–500). The right subplot compares these local weights (dotted lines) to the aggregated global weights (blue circles), which reflect a weighted average based on data volume. This decentralized training method allows data to remain local while still building a unified global model crucial for privacy-preserving collaborative learning

3.2.3 Differential Privacy Constraint

Figure 12 shows the trade-off between privacy (epsilon, ε) and noise (sigma, σ) in differential privacy. As ε increases (looser privacy), the required noise σ decreases, making the model more accurate.

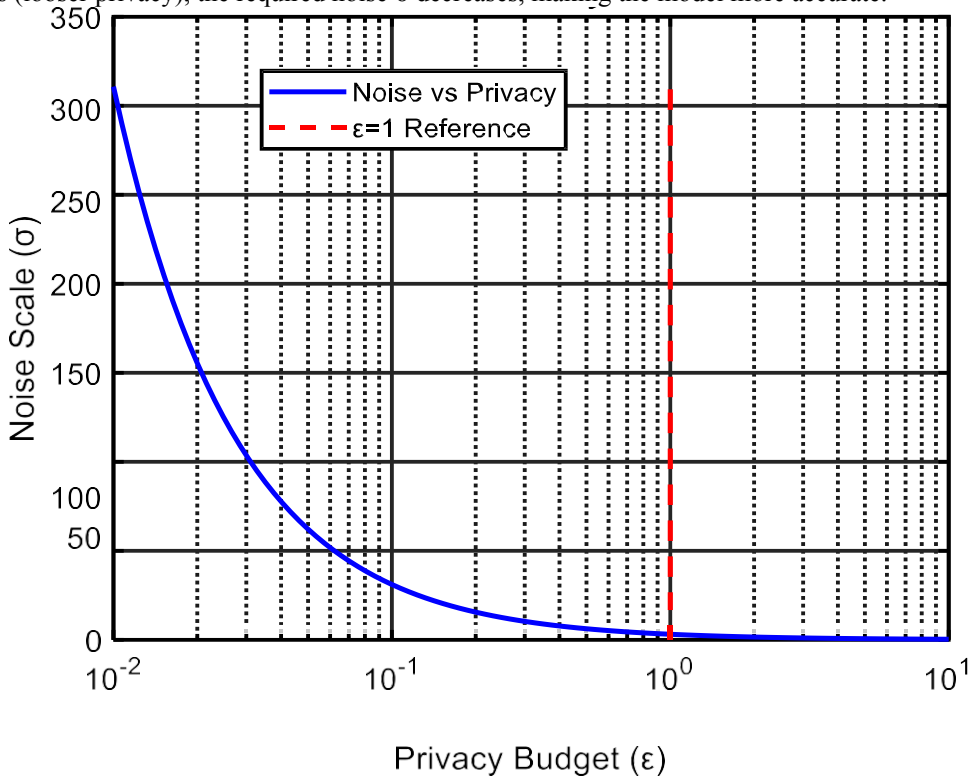

Figure 12: Differential Privacy Trade Off

At ε=1, marked with a red dashed line, noise still remains moderate. This curve emphasizes the balancing act: smaller ε improves privacy but adds more noise, potentially degrading performance. This graph provides insight into how privacy guarantees can be enforced mathematically while maintaining usable model accuracy.

IV. CONCLUSION

This paper addressed the critical challenge of balancing performance with privacy in secure access control systems for high-security environments. To this end, a novel framework integrating Federated Learning (FL) and Convolutional Neural Networks (CNNs) was conceived, developed, and evaluated.

One of the objectives was focused on designing a real-time facial recognition system using Convolutional Neural Networks, and it was successfully accomplished through the development of a deep-learning-based architecture capable of rapid feature extraction and classification. The CNN model demonstrated the ability to detect, track, and recognize faces with high precision, confirming that the system design was effective.

The second objective for this paper, involving the implementation of Federated Learning in the training process, was fully achieved by distributing model training across multiple clients without transferring raw facial data to a central server. This decentralized strategy ensured privacy preservation while enabling collaborative model improvement. The successful synchronization of client-side updates confirmed the feasibility of Federated Learning for sensitive biometric applications.

Contribution to knowledge:

i. **Integration of Federated Learning with CNN for Real-Time Facial Recognition:** This work introduces a novel framework that integrates Federated Learning with Convolutional Neural Networks to enable real-time facial recognition in high-traffic environments. By ensuring that model training occurs locally on edge devices without transferring raw facial data, the research contributes to the advancement of privacy-preserving AI in biometric systems.

ii. **Quantitative Evaluation of Privacy-Preserving Models in High-Density Scenarios:** This work provides empirical evidence showing that a Federated Learning-based CNN model can achieve high accuracy (**96.5%**), low latency (**150 ms**), and strong spoof detection while preserving user privacy. These results fill a critical knowledge gap in evaluating decentralized facial recognition systems under real-world, high-density conditions.

REFERENCE

[1]. Adeniyi, P. A. O. (2021). *Traffic mitigation and congestion in Ibadan, Oyo State Nigeria: Causes and solutions* (Doctoral dissertation). University of North Dakota. Retrieved from https://www.proquest.com/openview/aba3423a2bf990f2e53690597bb321ac/1?cbl=18750&diss=y&pq-origsite=gscholar

[2]. Adetoyi, O. E., & Awe, B. P. (2022). Face recognition enabled door access control system. *FUOYE Journal of Engineering and Technology, 7*(1), 28–31. https://doi.org/10.46792/fuoyejet.v7i1.656

[3]. Ahmad, J., Farman, H., & Jan, Z. (2018). Deep learning: Methods and applications. *Foundations and Trends in Signal Processing, 7(*3–4), 31–42. https://doi.org/10.1561/2000000039

[4]. Almeida, D., Shmarko, K., & Lomas, E. (2022). The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: A comparative analysis of US, EU, and UK regulatory frameworks. *AI and Ethics, 2(*3), 377–387. https://doi.org/10.1007/s43681-021-00077-w

[5]. Alsaleh, A. (2025). Toward a conceptual model to improve the user experience of a sustainable and secure intelligent transport system. *Acta Psychologica*, 255, 104892. https://doi.org/10.1016/j.actpsy.2025.104892

[6]. Anwarul, S., & Dahiya, S. (2020). A comprehensive review on face recognition methods and factors affecting facial recognition accuracy. *Imaging Science Journal, 68*(2), 1–15. https://doi.org/10.1080/13682199.2020.1738741

[7]. Arachchilage, S. W., &Izquierdo, E. (2019). A framework for real-time face-recognition. *2019 IEEE International Conference on Visual Communications and Image Processing (VCIP).* https://doi.org/10.1109/VCIP47243.2019.8965805

[8]. Ayomide, O. A., Dare, O., Olumoye, M. Y., Obiora Emeka, I., Kolawole, F., ... Virginus, N. (2021). Optimization of an identity access control system using biometric techniques. *International Journal of Progressive Sciences and Technologies, 27(*2), 647–653. Retrieved from http://ijpsat.ijsht-journals.org

[9]. Azis, A. D., Mamonto, A. A. N., Amiq, B., Rambe, K. M., &Syahputra, A. R. (2023). Facial recognition technology: A multinational analysis of regulatory framework, ethics, and legal implications in security and privacy. *International Journal of Science and Society, 5(*4), 498–510. https://doi.org/10.54783/ijsoc.v5i4.808

[10]. Babu, S. B. V., Manoranjini, J., Changala, R., Aarif, M., Mary, S. S. C., & Raj, I. I. (2024). Biometric-based access control systems with robust facial recognition in IoT environments. *2024 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS).*https://doi.org/10.1109/INCOS59338.2024.10527499

[11]. Baobaid, A., Meribout, M., Tiwari, V. K., & Pena, J. P. (2022). Hardware accelerators for real-time face recognition: A survey. *IEEE Access, 10,* 83723–83739. https://doi.org/10.1109/ACCESS.2022.3194915

[12]. Beitelmal, W. H., Nwokolo, S. C., Meyer, E. L., &Ahia, C. C. (2024). Exploring adaptation strategies to mitigate climate threats to transportation infrastructure in Nigeria: Lagos City, as a case study. *Climate, 12 (*8). https://doi.org/10.3390/cli12080117

[13]. Belidhe, S., Dasa, S. K., &Jaini, S. (2021). Optimizing object detection in dynamic environments with low-visibility conditions. *International Journal of Advanced Trends in Engineering and Technology, 6*(2), 2019–2022.

[14]. Betrand, C. U., Juliet Onyema, C., Eberechi Benson-Emenike, M., &Allswell Kelechi, D. (2023). Authentication system using biometric data for face recognition. *International Journal of Sustainable Development Research, 9*(4). https://doi.org/10.11648/j.ijsdr.20230904.12

[15]. Borkar, N. R., & Kuwelkar, S. (2017). Real-time implementation of face recognition system. *Proceedings of the International Conference on Computing Methodologies and Communication (ICCMC),* 249–255. https://doi.org/10.1109/ICCMC.2017.8282685

[16]. Carter, A. M. (2018). *Facing reality: The benefits and challenges of facial recognition for the NYPD*. State University of New York, Empire State College.

[17]. Chen, L., Kumar, A., and Zhang, J. (2023). Advancements in multi-modal biometric fusion: A review. IEEE Transactions on Biometrics, Behavior, and Identity Science, 5(2), 145-160. doi: 10.1109/TBIOM.2023.3245678

[18]. Chen, Z., Chen, J., Ding, G., & Huang, H. (2023). A lightweight CNN-based algorithm and implementation on embedded system for real-time face recognition. In *Multimedia Systems* 29(1), 129–138. https://doi.org/10.1007/s00530-022-00973-z

[19]. Deshmukh, S., Patwardhan, M., & Mahajan, A. (2016). Survey on real-time facial expression recognition techniques. *IET Biometrics, 5(*3), 155–163. https://doi.org/10.1049/iet-bmt.2014.0104

[20]. Dhaw, M. H., Zerek, A. R., Ali, N. K. A., Tohamei, W. M., & Emsalem, T. M. (2008). Biometric recognition system "Fingerprint and face recognition." *International Journal of Computer Science and Network Security, 8*(11), 1–10.

[21]. Du, H., Shi, H., Zeng, D., Zhang, X. P., & Mei, T. (2022). The elements of end-to-end deep face recognition: A survey of recent advances. *ACM Computing Surveys, 54(*10), 1–42. doi:10.1145/3507902

[22]. Elechi, P., Okowa, E., & Ekwueme, U. (2022). Facial recognition-based smart door lock system promise. *Journal of Scientific and Industrial Research, 6*(2), 95–105.

[23]. Elsherbiny, M. M., Radwan, A. M., Hussien, A. H., Salah, H. A., Galal, M. A., Saraya, S. F. (2024). Elevating smart home security and personalization through advanced face detection technology. *Communication and Computer Engineering Research Magazine, 1(*1), 1–13.

[24]. Erinjogunola, F. L., Sikhakhane-nwokediegwu, Z., Ajirotutu, R. O., &Olayiwola, R. K. (2025). Enhancing bridge safety through AI-driven predictive analytics. *International Journal of Social Science Exceptional Research, 4*(2), 10–26. https://doi.org/10.54660/IJSSER.2025.4.2.10-26

[25]. Essien, J. (2023). Enhancing role-based access control with embedded facial recognition RBAC-EFR system. *International Journal of Science and Research, 12(*6), 2767–2774. https://doi.org/10.21275/SR23625003927

[26]. Fadel, N. EL. (2025). Facial recognition algorithms: A systematic literature review. *Journal of Imaging, 11*(2). https://doi.org/10.3390/jimaging11020058

[27]. Fegade, V., Chodankar, A., Bhingle, A., & Mhatre, S. (2022). Residential security system based on facial recognition. *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, 803–811. https://doi.org/10.1109/ICOEI53556.2022.9776940

[28]. Feroze, S. A., Awan, S. ur R., & Ali, S. Z. (2024). The Facial Recognition Technology in Academic Attendance: A Comparative Study for Real-Time Management. *International Journal of Technology, Innovation and Management (IJTIM)*, *4*(1), 1–19. https://doi.org/10.54489/adxn2030