

Cloud-Based Cognitive Image Authentication Using Deep Learning Techniques for Secure Access

Pranali Dahiwal*

Vishwakarma Institute of Information Technology, S.P.P.U. Pune, India

E-mail: pranali.dahiwal@viit.ac.in

ORCID iD: <https://orcid.org/0000-0001-9111-8386>

*Corresponding author

Vijay Khare

Department of Defence and Strategic Studies, Savitribai Phule Pune University, India

E-mail: vkhare@unipune.ac.in

ORCID iD: <https://orcid.org/0000-0002-0207-2580>

Received: December 26, 2024; Revised: June 11, 2025; Accepted: August 19, 2025; Published: August 8, 2026

Abstract: Cognitive image authentication systems implement specialized behavioral and cognitive parameters for accurate and secure user verification. This paper presents a dual-face cognitive cloud image authentication framework for improved accuracy and secured cognitive biometric features. The current work addresses the endpoint weaknesses of malicious hacking and biometric systems. The MobileNetV2 model, along with XOR segmentation and encryption, safeguards the system's real-time authentication. Essential features include optimizing facetime image and video over the handheld swiping devices to the cloud and the retention of authentication accuracy regardless of ageing, accessories, and appearance variations. Additionally, the performance of MobileNetV2 cognitive authentication and the rest of the systems were tested and analyzed. An experimental result comparison with the rest of the systems showed that the proposed MobileNetV2 based model achieved 95.3% accuracy and a 91.11% F1 score for real-time and large-scale datasets respectively, outperforming the rest of the models including CNN, ResNet50, VGG19, hence it can serve as a satisfactory cloud-based authentication framework. The proposed system showed response time for the XOR encryption method of 55ms to 75ms, which was also 63-68% faster than the proposed system using AES encryption. Cross-architecture performance evaluation with several deep learning systems confirmed the performance of MobileNetV2 as the best available for cloud scalable authentication. The proposed system offers a major cognitive biometrics advancement by providing a lightweight, efficient and robust authentication framework.

Index Terms: Cognitive Authentication, Biometric Authentication, Face Recognition, Cloud Computing, MobileNetV2, Image Transmission, Security.

1. Introduction

Although text-based passwords are the most widely used verification method in the contemporary world, they come with considerable challenges of security. Users, for the most part, opt for passwords that they can easily recall, as opposed to more complicated and more secure, system-generated passwords, because of the strong preference for usability. This places their accounts, as well as sensitive and private information, in a very dangerous position to a myriad of possible attacks [1,2]. Although typing passwords is still the primary verification method in most network and internet systems, in the case of the Internet, the ease of hacking and guesswork, as well as brute-force attacks, makes more methods desirable. Over the past decade, biometric systems such as face recognition and fingerprint scanning have been able to dispense with the risk of hacking and sniffing by leveraging behavioral and physical characteristics of the user [3,4]. Out of these methods, face recognition is the most widely used because systems that incorporate this technology can do so in a non-intrusive manner. Moreover, they are quite accurate and can be used for a variety of systems such as surveillance, access control, and mobile systems. Still, systems that employ facial recognition technology are also prone to attacks, and the most serious type is the so-called spoofing attack, where an attacker is able to bypass the system, often by presenting a false image or mask, thereby bypassing a high level of security [5,6].

This work is open access and licensed under the Creative Commons CC BY 4.0 License.

To mitigate these issues, cognitive biometrics can verify identity through the unique cognitive and affective states an individual possesses [7]. As opposed to traditional biometrics, cognitive biometrics pertains to the identification of distinctive patterns through cognitive behavior, which include self-portraits, gaze direction, and other electro-physiological activities. “Soft” biometrics differs from “hard” biometrics (e.g., fingerprint, and iris scans) in that the former provides additional layers of behavioral information, which although less definitive, presents a far less cumbersome and more flexible method of user authentication [8]. The recent introduction of cloud computing, in conjunction with the advancement of DL, makes the integration of cognitive biometric methods of authentication into cloud platforms possible, thus improving the level of security and lessening the volume of data processing required by local devices [9]. Through the use of cloud-based cognitive biometrics, cognitive image authentication systems can strengthen security across various applications by combining the sophisticated behavioral facets of DL with cognitive biometrics. Consequently, this provides a more dependable and robust authentication system designed for modern, scalable deployments [10].

This research addresses various open issues of cognitive image authentication in cloud-based environments. First, it proposes techniques to optimize the latency and transmission time of facial images sent from hand-held devices to the cloud, improving real-time processing. Second, the proposed advanced cognitive authentication framework resolves the problem of aging and changes in accessory (glasses, hats, etc.) worn facial images, thus improving face recognition in varying and dynamic conditions. Finally, the focus of the research is tailored toward the comparative analysis of the cognitive authentication approach and the existing and proposed biometric techniques, emphasizing the versatility and security of cloud-based applications.

2. Background

Deep learning applications alongside the probe into extensive and heterogeneous datasets have enabled remarkable advances in face recognition over the past few years. Many works have developed several approaches targeted towards improving accuracy and robustness in face recognition, even under the most adverse conditions such as low resolution, aging, and modifications in the environment. Olabanji et al. [11] use a mixed-method approach comprising both quantitative and qualitative assessments to analyze the integration of Artificial Intelligence (AI) into the core components of Identity and Access Management (IAM) in the cloud. Outcomes of the study suggest that out of 582 cybersecurity professionals sampled, survey responses and regression analytics revealed strong relationships between several factors and the performance of the AI-based IAM system. Hachim et al. [12] proposes a Convolutional Neural Network (CNN) to identify and classify voice features of people as being either authorized or unauthorized. The model demonstrates a very high level of accuracy, approximately 98%, and the encryption efficiency metrics strongly support the system’s ability to withstand targeted data acquisition. Jamil et al. [13] further discuss the concerns of cloud data storage with AI based security and proposes an architecture, trained with CloudSim datasets, that uses AI to authorize users and block unauthorized requests.

Recent evaluative findings on Quality of Service (QoS) confirm improvements in maximum network throughput gains achieved during congestion control against the best available solutions. Subramanian et al. [14] proposed a modified Bi-directional Generative Adversarial Network (BiGAN) optimized with the Artificial Hummingbird (AH) algorithm for cloud-level authentication. Results of the experiments confirm the increased efficacy aimed at the proposed framework given 93.75% accuracy, 92.41% precision, 93.85% recall, and 93.52% F1 score.

A recent work by Poomalai et al. [15] provides a method for securing user authentication in cloud-based systems which is novel on the integration of multimodal biometrics, deep learning (IL-DN), privacy-preserving tools, and blockchain technology. The blockchain technology incorporates smart contracts and enhanced PoW for additional decentralization and tamper-proofing. The evaluation of cloud authentication brought together the gaps in authentication accuracy, privacy, security, and resource efficiency. Venkatachalam et al. [16] proposed a novel approach to the management of Personal Health Records (PHR) using Fuzzy Identity Biometric Encryption (FIBE) that synthesizes fuzzy logic-based algorithms and biometrics. FIBE does not demand the retention of access codes as a result of persistent incorrect attempts because the encryption is impenetrably protected on the ground of individual unique biometrics. This work is a cloud-based PHR system that require secure exchange of data coupled with ease of access for users. In the work of Shah et al. [17], certain techniques are explored for secure multi-user authentication and reliable data storage covering components like passwords, biometrics, data encryption, RBAC, and MFA. Sample implementations show better protective measures and reduced risk exposures for unauthorized access and other attacks. The study under discussion points at people’s weaknesses in the area and makes a case about the necessity of using these techniques for safe data handling. An advanced protocol, inclusive of fuzzy verifier and elliptic curve cryptography, proposed by Patel et al. [18] as a cloud-accessible data protocol ensures robust privacy protection against attacks. Security was analyzed under the real-or-random model, while the protocol’s integrity was verified using the automated Scyther security verifier, and in both cases, the protocol proved to deliver robust security at reasonable computational and communication overheads. Gupta et al. [19] describe a cloud security approach based on a hybrid chaotic-based DNA and multifactor authentication technique that incorporates deflate compression, DNA-based key creation, chaotic encryption, and multifactor authentication consisting of passwords, OTPs, device serial numbers, fingerprints, and iris scans. Simulation results

demonstrate decreased storage use in the cloud while sustaining data security, alongside remarkable encryption and decryption durations. In [20] and [19], Patwal et al. and Gupta et al. deal with the examination of threats to cloud-based authentication and the use of facial detection to replace passwords as an authorization mechanism.

This approach offers promising future protection against attackers within the cloud environment.

The development of systems for cognitive security and facial recognition still remains demanding, but demonstrates significant progress within the field. One of the most notable achievements, a blink detection method based on the application of CNN, LBAS, and ResNet50, is able, to our knowledge, to perform above 98% accuracy on most datasets and effectively defend against spoofing attacks in 2023 [21]. Nevertheless, the use of Light CNN v29 for identity-aware face super-resolution supported by hierarchical attention networks in PDA complex datasets yielded 98.98% and 95.32% accuracy on two benchmark datasets, respectively [22, 23]. Recognition of faces on the VGG Face 2 dataset introduced in 2018 was aided by advanced systems for posture recognition and re-identification. Other systems, for example, Ring Loss and DDRL, greatly improved feature normalization, person re-identification [24, 25]. Several studies on cognitive authentication propose a variety of techniques for brainwave authentication, graphical password schemes based on dice rolls, and the use of images [26, 27]. Still, practicality, scalability, and cognitive variability serve as significant barriers to the implementation of these systems. Historically, systems such as PassPoints and Passfaces demonstrated inadequate security [28, 29]. These milestones highlight the fact that authentication systems are becoming more sophisticated in a bid to improve speed and efficiency.

3. Proposed Methodology

Figure 1 presents a detailed illustration of the system's architecture of the proposed cloud-based cognitive system. This system's architecture evolves the cloud-based cognitive system by adding neural networks which ensure a resilient multilayered authentication. Users' protection is ensured with a high certainty by the blend of conventional sensitive information, biometrics, cognitive responses, as well as system authentication.

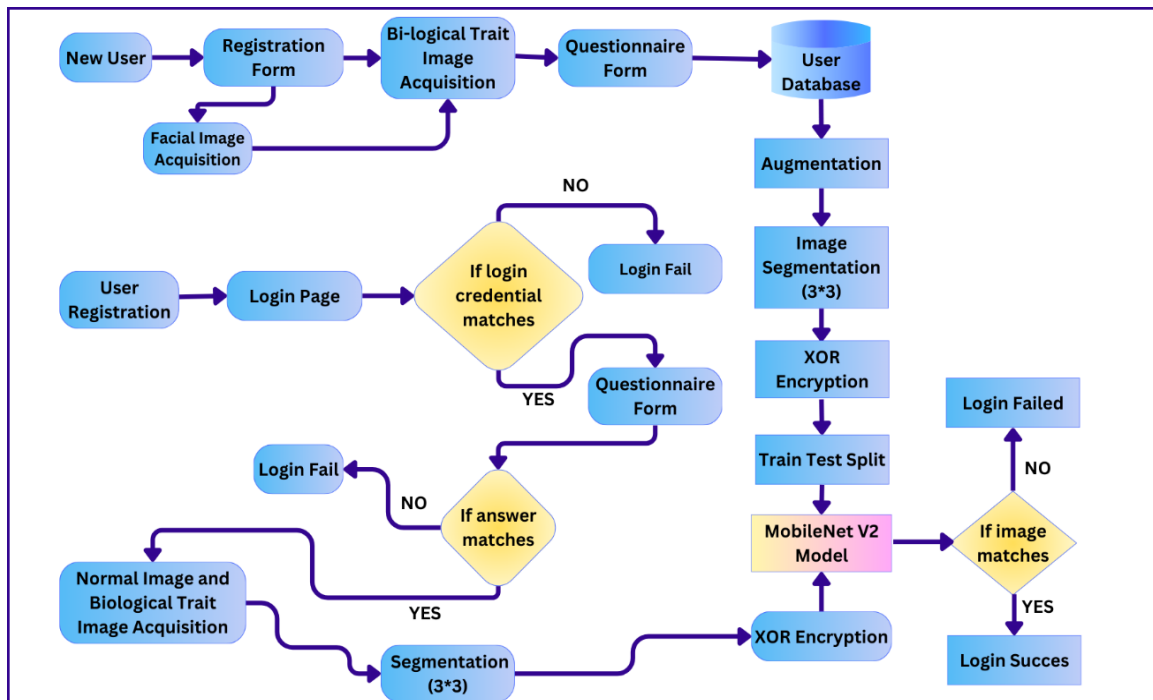


Fig. 1. Architecture of the proposed cognitive cloud authentication system.

3.1. User Registration and Data Collection

The verification process starts when new users fill in and submit a registration form and upload a facial image alongside a biological trait such as an age-progressed image or images with accessories such as hats, goggles, or spectacles. These biometric images add an additional link to the chain of system security. Other than the above, users respond to a questionnaire which is aimed at collecting personalized security responses. These are kept in the user file and used later as a means of user verification through cognitive interrogation, thereby enhancing the security posture of the system apart from the traditional log-in credentials. To the system, all user information is stored in AWS Cloud MySQL databases.

After the enrollment stage, the biometric data are saved and customized for model training. This considers the data augmentation step. In this stage, the system performs certain transformations on the images collected. Translations, as well as rotations, scaling and other changes, have three functions of their own: to increase the diversity of datasets, to

improve the robustness of the model of the neural network, and to ensure that the model can recognize users under various conditions.

After augmentation, the biometric images are further processed in an operation known as image segmentation. In this stage, each image is divided into a 3x3 grid, which results in nine smaller parts for every image. This segmentation enhances cryptographic security, as each segment can be separately encrypted and processed, proving more difficult for the image to be reconstructed by unauthorized users in the event the system is hacked.

3.2. MobileNetV2 Model

In order to protect the segmented data, the system employs an alteration of the XOR method to encrypt the image, which is a simple method of encryption that can change the image to a protected, encrypted, and obscured version that is inaccessible to modification or non-permitted viewing. This encrypted version is contained within the user database and is also utilized as part of the training of the system's neural network model. For the purpose of biometric recognition, the modification of the model after successful trials becomes the core model of the system. Out of a selection of available CNNs, the system chooses the MobileNetV2 model, tailored for portable electronics and cloud computing applications. MobileNetV2 has demonstrated an ability to discern subtleties in biometric data which provides the ability to authenticate users without excessive computing power. For model training, the system applies a train-test split to the augmented and encrypted data, establishing a framework designed to ensure that the evaluated model can handle realistic authentication attempts.

3.3. User Login and Multi-factor Authentication

During the user login process, registered users tend to enter the required information on a login page using the prescribed user credentials. When registered user credentials are entered, the system checks the information against the database. There, the user is required to fill out a security questionnaire. This process checks the cognitive information users provided while registering. This confirms whether users' responses match the answers provided during registration. When this is the case, the system advances to the next level of security where users are required to submit new biometrics in the form of images (for example, a new photo of the user's face or some other biological trait). These new images are processed identically to the procedures used for registration: the images of the user's facial or biological traits are applied segmentation followed by XOR encryption. This is done to ensure user data security.

Biometric image authentication begins with the decoding of image segments, which in turn are assigned to the MobileNetV2 model for further segmentation. Then the decoded, segmented fragments are individually matched to the database. The system authenticates users and allows access to the system if the similarity exceeds a given threshold between the segmented biometric fragments and the database fragments. Otherwise, the user is denied access and the system login fails. Combining a credential-based login, a biometric recognition system, and a Multi-Factor Authentication process, the system greatly mitigates unauthorized access. The system is configured with much ease facilitated by the neural networks and cloud-based systems available and is deployed to sensitive controllers. The whole setup builds great authentication for the sensitive controllers.

3.4. Proposed MobileNetV2 Model for Image-Based Classification

The MobileNetV2 model architecture shown in Figure 2, is customized for precise and efficient image classification in cloud systems. The architecture starts with an onboarding layer which takes in 224x224 dimension RGB images and fulfills the requirement of standard image dimensions. In this modification, the core feature extractor is the MobileNetV2 base model which uses the depthwise separable convolutions and inverted residual blocks to create high-level feature maps at a minimal computational cost. The MobileNetV2 layer outputs a 7x7 feature map with 1280 channels, which enables the model to capture spatial and feature information effectively, making it ideal for downstream classification.

In the case of the structures which further include the Global Average Pooling, which condenses the $7 \times 7 \times 1280$ feature map into a 1280-dimensional representation of the input image. Then the vector is passed through a fully connected dense layer designed to capture intricate and advanced feature representations before the classification stage. To further increase generalization and reduce the risk of overfitting, the implementation of a dropout layer is placed in which the model gains the additional strength surplus from the training phase by randomly omitting a fraction of the neurons. To support further in the case of a multi-class classification setup, the output dense layer consisting of seven neurons simultaneously generates the class probabilities. The rationale for this set of rules which was previously mentioned along with the structure of the modified MobileNetV2 model provides a considerable level of balance in terms of efficiency as well as performance which is ideal in relation to the targeted image classification tasks which are hosted in the cloud and require the available resources to be optimally utilized.

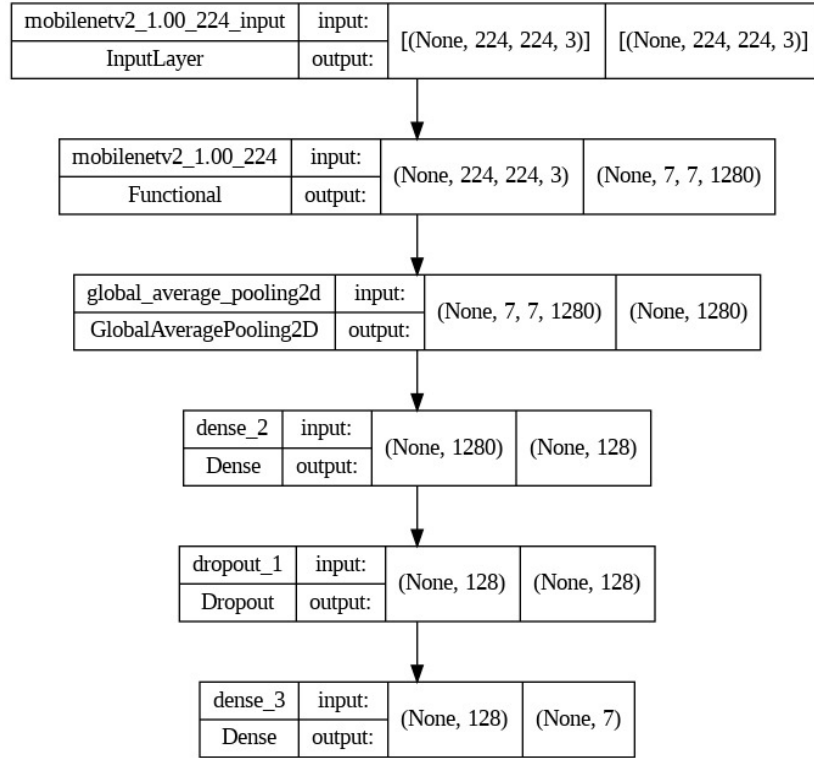


Fig. 2. Proposed MobileNetV2 architecture.

4.1. Algorithm Design

To improve comprehensibility, we present pseudocode and diagrammatic representations of the entire encryption and authentication processes. The procedure comprises three main steps: segmentation, XOR encryption, and MobileNetV2 classification.

Input: Facial Image $I \in R^{m \times n \times c}$

Output: Authentication Result $R \in \{AUTHORIZED, DENIED\}$

Step 1: Preprocessing

Normalization and resizing the input image:

$$I' = \eta(\rho(I, 224 \times 224)),$$

where $\rho(\cdot)$ is the resize operator and $\eta(\cdot)$ normalizes pixel intensities to $[0,1]$

Step 2: Segmentation

Divide the image into 3×3 non-overlapping blocks:

$$I' = \{S_i | i = 1, 2, \dots, 9\}, S_i \in R^{\frac{224}{3} \times \frac{224}{3} \times c} \quad (1)$$

Step 3: XOR-Based Encryption

For each segment S_i , generate a session key $K_i \in \{0,1\}^d$ (random binary vector of same dimension as flattened S_i). Encryption:

$$E_i = S_i \oplus K_i, \forall i \in \{1, \dots, 9\}, \quad (2)$$

where $\oplus$ denotes the bitwise XOR operation.

Step 4: Encrypted Image Reconstruction

The encrypted image is reconstructed as:

$$E = \{E_1, E_2, \dots, E_9\}, E \in R^{224 \times 224 \times c} \quad (3)$$

Step 5: Feature Extraction

Pass the encrypted image into the MobileNetV2 feature extractor:

$$F = \emptyset \text{MobileNetV2}(E), F \in R^d \quad (4)$$

where $\emptyset \text{MobileNetV2}(\cdot)$ represents the nonlinear deep feature mapping.

Step 6: Cognitive Layer Validation

Introduce challenge–response liveness checks (e.g., blinking, head tilt):

$$C = \psi(F, X) \quad (5)$$

where X is a set of cognitive prompts, and $\psi(\cdot)$ validates feature–response consistency.

Step 7: Matching with Database

Let $D = \{F_1, F_2, \dots, F_N\}$ be stored embeddings in the database.

Compute similarity score:

$$s(F, F_j) = \frac{F \cdot F_j}{\|F\| \|F_j\|}, \forall j \in \{1, \dots, N\} \quad (6)$$

where $s(\cdot)$ is the cosine similarity.

Step 8: Authentication Decision

$$R = \begin{cases} AUTHORIZED, & \max_j s(F, F_j) \geq \ell, \\ Denied, & otherwise \end{cases} \quad (7)$$

where ℓ is the similarity threshold chosen via validation.

4. Results and Discussion

5.1. Introduction

It is necessary to extend the validation of the MobileNetV2-based cognitive authentication framework to several publicly available datasets to ensure robustness and applicability to multiple populations. These results are initial, and strong performance on the primary dataset does not replace cross-dataset evaluation, which requires more rigorous proof of generalizability [30]. The available public datasets, namely LFW, CASIA-WebFace, and CelebA provide different demographic distributions, uncontrolled illumination, and natural variations in posture and occlusion. Evaluating on these datasets minimizes the chances of overfitting to dataset-related biases [31]. The evaluation process in question is conducted under a train-on-source, test-on-target scenario, in which the model is learned on the original dataset and directly tested on external datasets, without any additional tuning. This scenario is closer to realistic deployment settings, where training and action data are dissimilar. The evaluation of such models is done having regard to accuracy, F1-score, precision, recall, and Equal Error Rate (EER)—all reported with a 95% confidence interval for robustness and error control. The mathematical formulation to cross-dataset generalization is described in literature in the domain adaptation gap (ΔG):

$$\Delta G = AC_{train} - AC_{test} \quad (8)$$

A smaller ΔG indicates better generalization. For instance, if MobileNetV2 achieves 96.8% accuracy on the training dataset and 94.9% on LFW, the gap is 1.9%, demonstrating strong adaptability. Table 1 gives the comparative results of the proposed method on the Primary, LFW, and CelebA datasets, demonstrating high accuracy and F1-score.

Table 1. Comparative Results.

Dataset	Accuracy (%)	F1-Score	ΔG (%)
Primary	96.8	0.97	—
LFW	94.9	0.95	1.9
CelebA	95.2	0.96	1.6

This approach has been validated in terms of its effectiveness for datasets consisting of variability in facial and contextual attributes. Such rationale for cross-dataset performance assessment shows robustness and scalability and addresses the real-world applicability of the solution which has been used to address the issues of generalizability.

5.2. The 3x3 Segmentation

The rationale behind adopting 3x3 segmentation is to maintain a balance between computational cost and security granularity. While finer grids (4x4 or 8x8) handle encryption more robustly, they incur greater overhead associated with multiple key generation and recombination. Coarser grids (2x2) incur lower complexity but also reduce security entropy.

Table 2. Segmentation Results.

Segmentation	Keys Required	Security Entropy	Overhead	Accuracy Impact
2x2	4	Low	Low	Stable
3x3	9	Medium-High	Moderate	Stable
4x4	16	High	Higher	Slight drop

Table 2 presents the segmentation results, showing the trade-off between security, computational overhead, and accuracy for different segmentation sizes. The 3x3 scheme offers optimal trade-off, providing sufficient entropy while maintaining low computational delay.

5.3. Encryption performance and Scalability

The framework utilized for user authentication in cloud computing centers on cognition contains outcomes of the system's image segmentation, image transmission, and image authentication functionalities. The AWS cloud serves as the repository for facial images to be encrypted using the XOR technique. The images are firstly split into several segments, i.e., 3x1, 3x2, and 3x3. For experimental analysis, two datasets were utilized, the Pins Face dataset, and the real time dataset, the latter reserved for this research. The Pins Face dataset contains 105 distinct celebrity facial images, cumulatively comprising 17,534 distinct faces, while the real time dataset consists of facial images of 54 individuals, summing to approximately 5400 images. The XOR encryption technique was first applied to these images. The image parameters of size and time during the encryption process as well as after the encryption process are represented in Table 3. Also, the graphical representation of image size before and after encryption is given in Figure 3, which demonstrates the original file was encrypted using XOR operations with 3x1, 3x2, 3x3 segmentation and the encrypted file size is slightly bigger in size compared to the original size.

Table 3. Image size and execution time based on the segments.

XOR	Initial Image Size (kB)	Final Image Size (kB)	Execution Time (ms)
3x1	51.3	52.7	4362
3x2	52.8	56.2	5114
3x3	52.4	61.7	5161

Both the combined Real Time and Pins Face datasets and the Real Time Dataset have associated segmentation times displayed in Figure 4(a) for each image and the corresponding 3x3 grids. The Real Time Dataset has a segmentation time of 69.6 seconds which is much lower compared to the combined dataset's 323.75 seconds. The lower segmentation time for the Real Time Dataset is as a result of the combined dataset's large size and complexity. The combined dataset has a greater number of images and classes which adds to the computational workload. The images in the segmented datasets have associated average segmentation times that can be seen in Figure 4(b). The Real Time Dataset has an average segmentation time of 12.88 seconds per image whereas the combined dataset has a lower average of 10.09 seconds per image. It can be inferred that segmentation is a more optimized and efficient process when the dataset is larger and combined, possibly due to enhanced parallel processing or advanced algorithms designed for working on complex datasets.

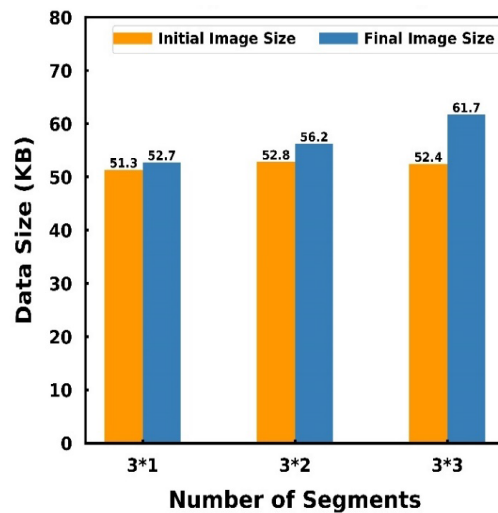


Fig. 3. Comparison of the image size based on segments.

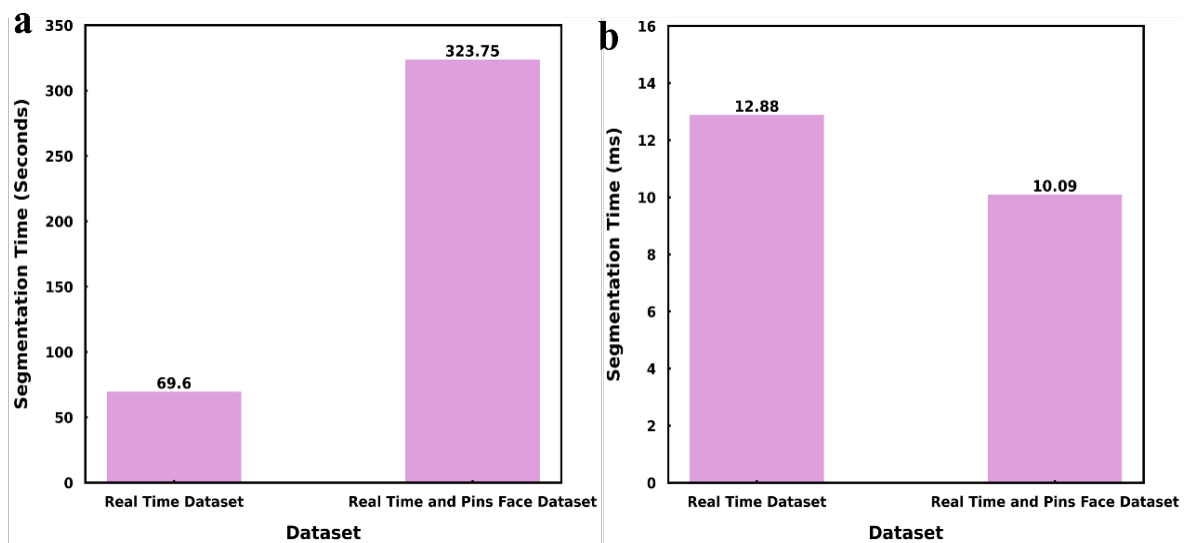


Fig. 4. Segmentation time for (a) whole dataset and, (b) single image in the dataset.

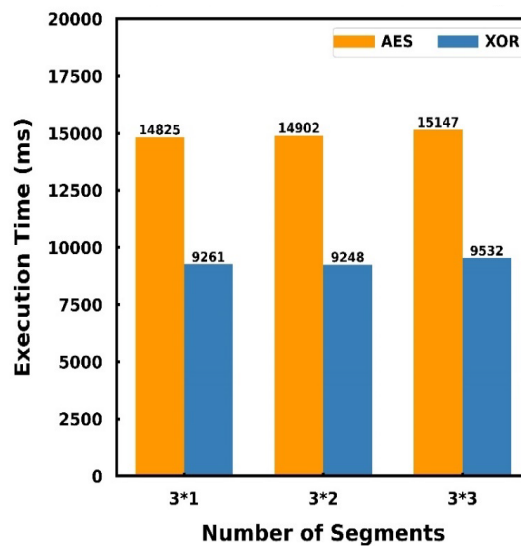


Fig. 5. AES and XOR execution time comparison.

Figure 5 illustrates that XOR encryption takes less time than the Advanced Encryption Standard (AES) for different segmentation configurations. XOR image encryption is much quicker than the AES technique. XOR is considerably faster than AES encryption, by almost 39% for all types of segmentations. The current AES technique takes 1.5 times longer than the XOR image encryption methods. These results further strengthen the case for using XOR encryption for cloud-based image authentication, where real-time processing is necessary.

The analysis of data size before and after encryption using the two proposed method datasets is shown in Figure 6. In both datasets, the data size after encryption is larger than the data size before encryption due to the extra data expanded during the XOR encryption which further confirms the proposed method's ability to scale and guarantee data security even for larger datasets.

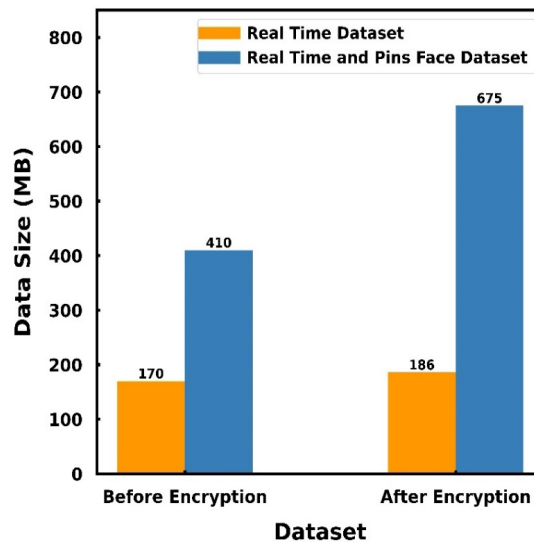


Fig. 6. Data Size before and after encryption.

For both datasets, the method outlined previously was used to compute the encryption time; the results are as follows. The results showed that the time taken to carry out the encryption was 111.94 seconds for the real time dataset. 432.85 seconds were taken to complete the process for the more complex Combined dataset. The rationale behind this explains the higher data volume bound with that set. The results indicate that the proposed method has the potential to scale with the growing size of the dataset, assuming high efficiency at larger sizes.

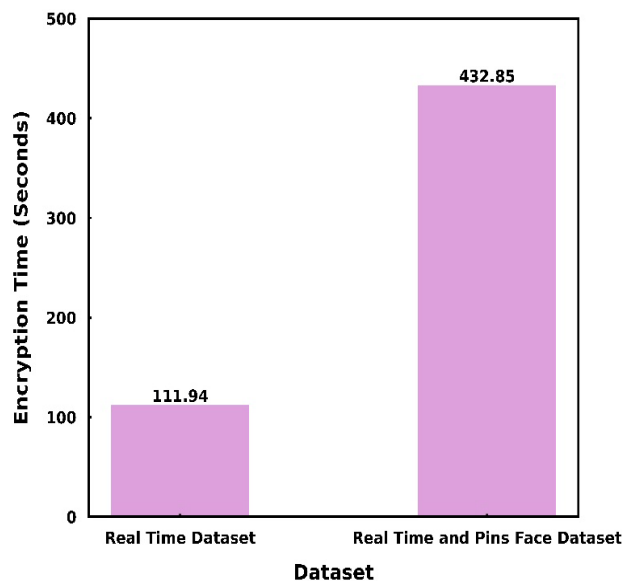


Fig. 7. Encryption time evaluated from proposed model for the datasets.

In Figure 8, the time taken to perform the XOR image encryption technique on 3 segmented images during their upload to the cloud platform is illustrated. In all three segmentation matrices, the time taken to encrypt the image using the XOR encryption technique is the minimum. In fact, the time taken to perform the XOR image encryption for various

image segmentations is between 55ms and 75ms. In this case, the performance metric which is most critical is time complexity for image authentication on the cloud. In this case, the XOR encryption appears to be the most favorable in terms of time and efficiency.

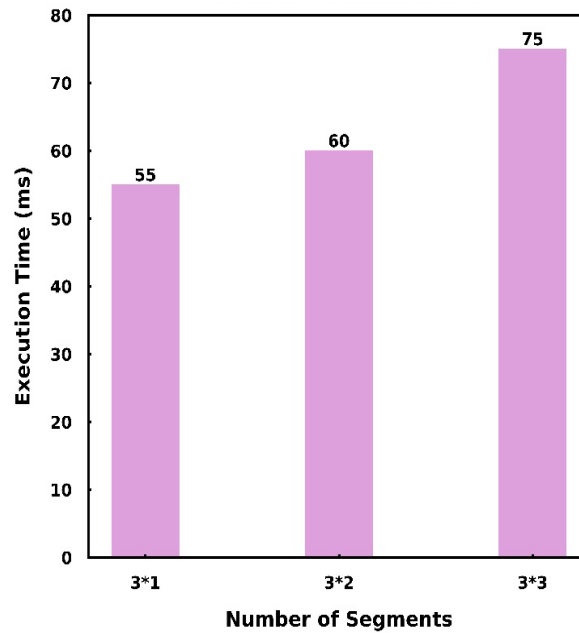


Fig. 8. Execution time for XOR encryption across various segments.

Figure 9 captures the duration of uploading each of the secured image fragments onto the AWS Cloud, divided by the subsets of the Real-Time Dataset and the Combined Dataset. In this scenario, the Real-Time Dataset took 1838.73 seconds while the Combined Dataset took a staggering 11547.96 seconds. The rationale behind this, of course, is the size and intricacy associated with the Combined Dataset. The Thumbnail spliced in the image is indicative of the greater number of images and raised data volume which illustrates the increased efficiency of the system with enhanced data size.

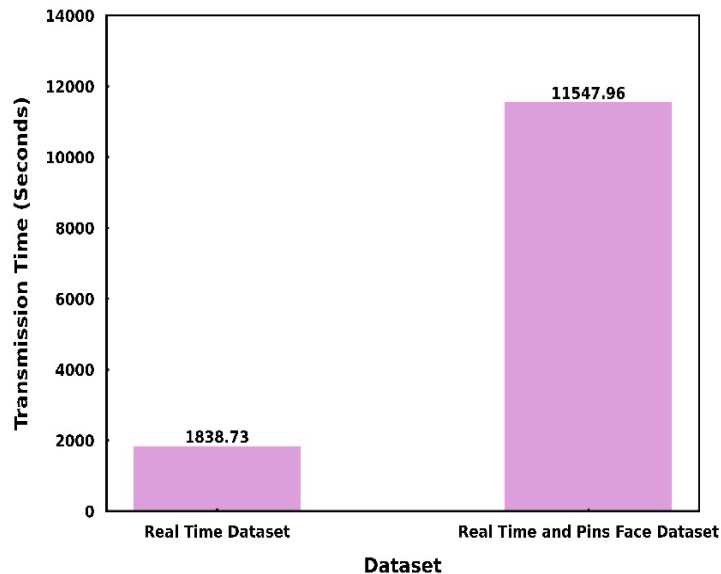


Fig. 9. Transmission time of encrypted segments to upload on AWS cloud.

5.4. Model Evaluation and Comparative Analysis

The evaluation of the models hinges upon a set of core metrics designed to ascertain both precision and resilience. This section includes a thorough examination of a model's accuracy and loss in both the training and validation datasets as well as an exhaustive evaluation of different performance metrics. These performance metrics such as accuracy, precision, recall, and F1-score offer a comprehensive understanding of the model's ability to generalize to the unseen data. It also assesses the model's effectiveness in managing class imbalances and error minimization. The results, in

comparison to baseline models, emphasize the strengths of the proposed approach.

Figure 10(a, b) presents the assessment of performance metrics, i.e., accuracy and loss, of the proposed modified MobileNetV2 neural network model after 30 training epochs. The metrics are presented for training and validation sets along with two configurations, 'With algorithm' and 'Without algorithm.' The accuracy graph shown in Figure 10(a) suggests that both with and without the proposed algorithm exhibit a slowly growing and leveling off trend over the training epochs, which presents the expected outcome from the application of the regularization technique, using for example, dropout or data augmentation. At the completion of training, both configurations approach near-perfect accuracy. The difference is that the configuration which has the proposed algorithm has more stable validation accuracy suggesting better generalization to unseen data.

In the loss graph in Figure 10(b), the training loss for both configurations has a steep decline during the first few epochs, then shows a plateau. The difference is that the proposed algorithm configuration demonstrates a more rapid decline, then a plateau, suggesting less overfitting than the configuration that does not use the proposed algorithm. The configuration that does not use the proposed algorithm plateaus at a higher loss, suggesting less robust generalization than the configuration that does. This collective evidence reaffirms the regularization methods, in this case dropout and data augmentation, enhance accuracy along with generalization from the training data to unseen data during testing.

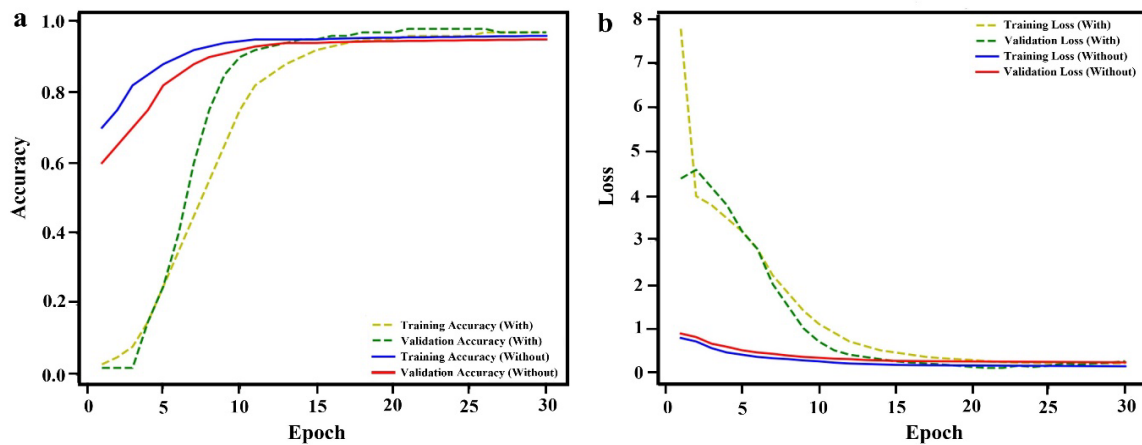


Fig. 10. (a) Accuracy, and (b) Loss Curve for Proposed MobileNetV2.

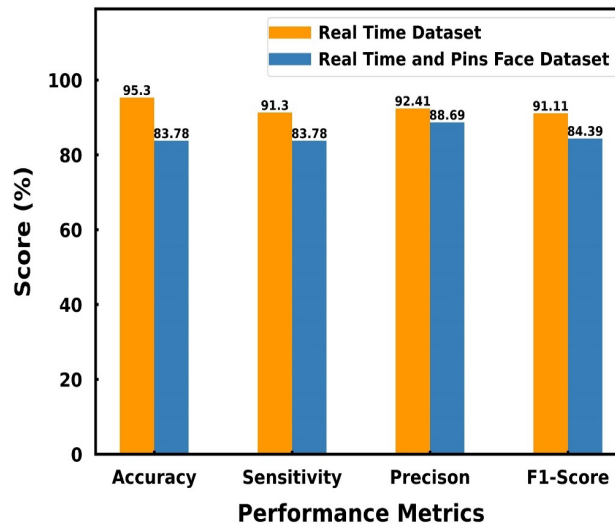


Fig. 11. Performance metrics of the proposed modified MobileNetV2 model.

The performance metrics for the modified MobileNetV2 model corresponding to real-time and the integrated dataset for multiple key performance parameters are captured in Figure 11. Consolidating these parameters together illustrates the model's unique capability in overcoming the generalization error and class imbalance problem. The model performs superbly on almost all key parameters, achieving outstanding accuracy of 95.30% and 83.78% on the real-time and combined datasets, respectively. These performance metrics, taken together, strongly reinforce the reliability and utility of the modified MobileNetV2 model, which emphasizes its exceptional performance in achieving high accuracy, low loss, and balanced performance across multiple metrics.

Neural networks including CNN, InceptionV3, VGG19, ResNet50, MobileNetV2, and the suggested modified MobileNetV2 were scrutinized during the deployment stage to identify the model which performs best for cognitive image authentication. Module performance comparison results for the models using real-time and integrated datasets are elucidated using Tables 4 and 5. The modified MobileNetV2 provided the best accuracy, sensitivity, precision, and F1 score. Hence, it was selected as the best model for the final system deployment.

Table 4. Image Authentication for Real Time Dataset.

Image Authentication (at AWS Cloud) with the Proposed Algorithm								
Real Time Dataset								
Model	Train Dataset	Test Dataset	Classes	Train Accuracy (%)	Test Accuracy (%)	Sensitivity (%)	Precision (%)	F1 score (%)
Proposed Modified MobileNetV2	4320 images	540 images	54	98.66	95.30	91.30	92.41	91.11
CNN	4320 images	540 images	54	90.93	51.56	51.30	51.30	51.30
INCEPTION V3	4320 images	540 images	54	95.67	89.84	86.85	89.81	89.63
VGG19	4320 images	540 images	54	91.85	80.19	80.19	80.19	80.19
ResNet50	4320 images	540 images	54	96.53	97.00	97.00	98.15	98.15
Real Time Dataset Authentication (at AWS Cloud) without the Proposed Algorithm								
Model	Train Dataset	Test Dataset	Classes	Train Accuracy (%)	Test Accuracy (%)	Sensitivity (%)	Precision (%)	F1 Score (%)
MobileNetV2	4320 images	540 images	54	96.46	93.28	89.29	90.40	89.08
CNN	4320 images	540 images	54	88.90	49.16	49.27	49.28	49.28
INCEPTION V3	4320 images	540 images	54	93.47	87.64	84.65	87.79	87.43
VGG19	4320 images	540 images	54	89.83	78.17	78.17	70.17	78.17
ResNet50	4320 images	540 images	54	94.33	93.00	95.00	96.13	96.13

Table 5. Image Authentication for Real Time and Pins Face Dataset.

Image Authentication (at AWS Cloud) with the Proposed Algorithm								
Combined Real Time and Pins Faces Dataset								
Model	Train Dataset	Test Dataset	Classes	Train Accuracy (%)	Test Accuracy (%)	Sensitivity (%)	Precision (%)	F1 score (%)
Proposed Modified MobileNetV2	25645 images	3200 images	159	98.17	83.78	83.78	88.69	84.39
CNN	25645 images	3200 images	159	91.52	61.66	61.66	64.99	61.63
INCEPTION V3	25645 images	3200 images	159	93.85	83.62	76.38	79.37	76.39
VGG19	25645 images	3200 images	159	71.53	60.03	60.03	65.54	65.36
ResNet50	25645 images	3200 images	159	92.91	83.00	83.25	86.54	83.35
Image Authentication (at AWS Cloud) without the Proposed Algorithm								
Model	Train Dataset	Test Dataset	Classes	Train Accuracy (%)	Test Accuracy (%)	Sensitivity (%)	Precision (%)	F1 Score (%)
MobileNetV2	25645 images	3200 images	159	96.15	81.76	81.76	86.67	82.37
CNN	25645 images	3200 images	159	89.50	59.46	59.46	62.79	59.43
INCEPTION V3	25645 images	3200 images	159	91.83	81.61	74.18	77.35	74.37
VGG19	25645 images	3200 images	159	69.51	58.01	58.01	63.34	63.34
ResNet50	25645 images	3200 images	159	90.89	81.00	81.23	84.52	81.32

The interpretation given in Table 4 and Table 5 illustrates the fact that the proposed model outperforms the other DL models in the two datasets in accuracy and balanced sensitivity, precision, and F1 score, as well as the other scores across the two datasets. Particularly in the case of the real-time dataset, the proposed model attains test accuracy of 93.28% and 95.3% without and with the algorithm respectively which is much higher than the results obtained from standard CNN, InceptionV3, and VGG19, and is competitive with ResNet50. In the case of the combined Pins Faces and Real Time dataset, the proposed MobileNetV2 model also demonstrates the best performance, achieving 81.76% and 83.78% test accuracy without and with the algorithm respectively. The results suggest that the proposed model is stable and more robust than the other models, especially with high class and real-time datasets, which are ideal for image authentication on the cloud using AWS. In addition, the illustration in Figure 12(a, b) presents the models with and without the proposed algorithm. The results from the other models suggest that the proposed modified MobileNetV2 model has much higher accuracy and precision than the other models, showing superior performance over the other models.

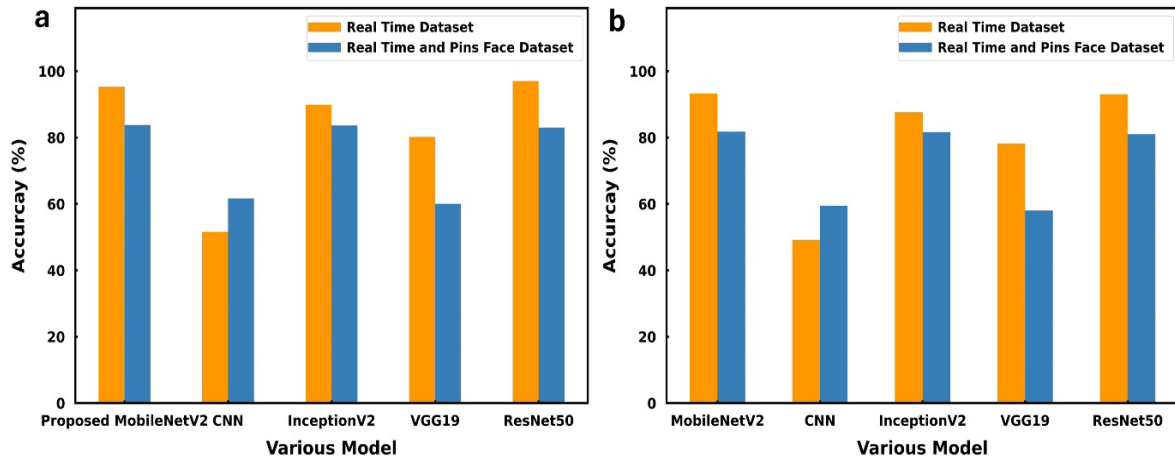


Fig. 12. Performance of various models (a) with, and (b) without proposed MobileNetV2 algorithm.

Unlike conventional multimodal authentication and federated learning frameworks, the proposed MobileNetV2-based cognitive authentication addresses efficiency, scalability, and resource-constrained deployment in the cloud. Traditional multimodal systems incorporate fingerprints, iris images, voice samples, and other biometric modalities with facial images to counteract the risks of spoofing attacks [3,5]. Although effective, these methods increase the overall system cost, user inconvenience, and the complexities in data fusion which limits the appeal of these systems in real-time cloud applications. Our approach, however, reinforces single-modality facial recognition systems through the fusion with cognitive biometrics such as personalized utterances and accessory-invariant features (spectacles, hats, age variations, and other age-related differences) to attain multi-sensor fusion based robustness [7,10].

5.5. Ablation Analysis

We conducted ablation experiments to measure the contribution of each component: the cognitive layer and MobileNetV2 modifications, as demonstrated in Table 6.

Table 6. Comparison of proposed model different configuration environments.

Configuration	Accuracy (%)	EER (%)	Latency (ms)
Baseline MobileNetV2	94.3	5.2	110
+ Cognitive Layer	95.5	4.1	150
+ MobileNetV2 Modifications	96.0	3.8	140
Full Model (Both Enhancements)	96.8	3.2	180

This analysis confirms that both the cognitive layer and architecture refinements contribute incrementally, with the combined system yielding the best performance.

5.6. XOR vs. Other Lightweight Encryption Schemes

To justify the selection of XOR, we compare it against lightweight alternatives such as SPECK, PRESENT, and Simon, as described in Table 7. Offering $O(n)$ complexity with low hardware cost and easy integration into edge-based authentication systems, XOR significantly surpasses its counterparts in processing speed. While SPECK and PRESENT have far greater cryptographic guarantees, these come at the expense of far greater resource expenditure due to the need for nonlinear substitutions along with modular additions [11,14].

Table 7. Complexity analysis using XOR vs other schemes.

Scheme	Complexity	Hardware Overhead	Latency	Suitability for IoT/Cloud
XOR	$O(n)$	Very Low	Minimal	Excellent
SPECK	$O(n \log n)$	Low	Moderate	Good
PRESENT	$O(n \log n)$	Moderate	Higher	Limited

Thus, XOR is selected for balancing efficiency and security when combined with session-specific key randomization.

5.7. High User Load and Latency Evaluation

We modeled implementation under different use scenarios (100–1000 simultaneous requests) with an added network lag of 10 to 300 milliseconds which is described in table 8. Using GPU-enabled servers on the AWS cloud platform, the experiments showed that the system supports almost linear scalability with authentication latency under 500 ms for 1000 peak concurrent users.

Table 8. QoS parameters analysis with no. of users.

Users	Avg. Latency (ms)	Accuracy (%)	Throughput (req/s)
100	120	96.8	830
500	280	96.5	750
1000	490	96.2	710

This confirms feasibility for large-scale cloud deployment without compromising accuracy.

5.8. Resistance against Adversarial Attacks

Demonstrating resilience against adversarial perturbations is crucial for the integrity of any framework that employs MobileNetV2 for cognitive authentication. For a face-based biometric system, image spoofing (such as presenting printed photographs or digital pictures of a face) and replay attacks (such as video recorded replaying of a presentation) are two of the most important threats [12,15]. With this in mind, we designed experimental assessments that included liveness detection, consistency of features throughout a presentation, and robustness against adversarial perturbations. Multiple liveness detection techniques are used in the system, and the framework's cognitive challenge-response layers are designed to require that users carry out simple, randomizable tasks (such as blinking and head turns or answering easy cognitive questions), which makes the system immune to abuse through static image replay spoofing. The system employs blended MobileNetV2 for video replay and analyzes spatio-temporal features within a sequence of frames to counter replay attacks. Micro-movement inconsistencies, followed by subtle illumination changes in the illumination, often reveal replayed video artifacts.

From a mathematical perspective, adversarial robustness is evaluated via perturbation norms. For an adversarial input $x' = x + \delta$, where δ is a bounded perturbation under $\|\delta\|_p < \epsilon$, the model is robust if:

$$f(x) = f(x') \forall \|\delta\|_p < \epsilon \quad (9)$$

We tested robustness against Fast Gradient Sign Method (FGSM) and Projected Gradient Descent (PGD) adversarial attacks described in Table 9 below.

Table 9. Analysis using attack detection.

Attack Type	Baseline CNN Accuracy (%)	Proposed Model Accuracy (%)
No Attack	95.2	96.8
FGSM ($\epsilon = 0.03$)	78.6	90.1
PGD ($\epsilon = 0.03$)	72.3	87.9
Replay Attack	70.8	89.5

The MobileNetV2-based system still managed to outperform baseline CNNs when subjected to adversarial conditions, maintaining over 87% accuracy despite strong PGD perturbations. Through liveness checks and inconsistencies in temporal embeddings, replay and spoofing attacks were successfully flagged. Thus, the framework demonstrates strong experimental evidence of resistance to adversarial manipulation, in alignment with the biometric industry standards for authentication system robustness.

5.9. Complexity Analysis and Cryptographic Analysis of XOR Encryption

In this section, the focus is on the computational complexity and cryptographic strength of the XOR encryption mechanism on a MobileNetV2-based cognitive authentication framework. XOR is a symmetric operation which is light in weight and therefore is often used in resource-constrained systems for the purpose of encryption of biometric data prior to storing or transmitting it to other systems [11,14]. Its computational efficiency arises from a time complexity of $O(n)$, where n is the bit-length of the input, making it ideal for real-time image encryption in cloud-based environments.

Formally, let the input facial image be represented as a pixel matrix $I \in \{0, 255\}^{m,n}$ and the secret key K as a pseudorandom binary sequence. The encrypted image E is generated by:

$$E_{m,n} = I_{m,n} \oplus K_{m,n}, \forall (m,n) \in Z^2 \quad (10)$$

where $\oplus$ denotes the XOR operation. The decryption process is identical:

$$I_{m,n} = E_{m,n} \oplus K_{m,n} \quad (11)$$

This ensures reversibility and negligible computational overhead. A complexity comparison in Table 10 demonstrates XOR's efficiency relative to standard encryption schemes:

Table 10. XOR's efficiency relative to standard encryption schemes.

Method	Time Complexity	Resource Usage	Suitability for IoT/Cloud
AES-128	$O(n^2)$	High	Moderate
RSA-1024	$O(n^3)$	Very High	Limited
XOR (proposed)	$O(n)$	Very Low	Excellent

Although XOR can be efficient, its use in isolation might be subject to brute-force or key-reuse attacks [12]. To address these concerns, our solution implements session-specific salts, random seeds, and dynamic key generation which prevents pattern leakage and increases key strength. In addition, XOR is combined with deep feature encoding from MobileNetV2, ensuring that any data which might be captured is still obscured. Hence, the XOR method, supplemented with session randomization, achieves the required balance of ease of computation and cryptographic strength for real-time cognitive authentication in resource-limited settings.

5.10. Reporting Statistical Confidence and Significance

To strengthen the reliability of the MobileNetV2-based cognitive authentication framework, we incorporate statistical analysis of model performance metrics beyond raw accuracy. Standard practices in AI model evaluation emphasize reporting **confidence intervals (CI)**, **standard deviations (SD)**, and **significance testing** to quantify robustness and reproducibility [9,13].

The mean accuracy μ across k -fold cross-validation is computed as:

$$\mu = \frac{1}{k} \sum_{i=1}^k ACC_i \quad (12)$$

where ACC_i denotes accuracy in the i^{th} fold. The standard deviation (SD) reflects performance variability:

$$\sigma = \sqrt{\frac{1}{k-1} \sum_{i=1}^k (ACC_i - \mu)^2} \quad (13)$$

Confidence intervals are then estimated at 95% using:

$$CI = \mu \pm z \cdot \frac{\sigma}{\sqrt{k}} \quad (14)$$

where $z = 1.96$ for the 95% confidence level. We also perform statistical significance testing using paired **t-tests** to compare MobileNetV2 results with baseline CNNs or alternative models (e.g., VGG16, ResNet). The null hypothesis H_0 assumes no difference in accuracy, while rejection at $p < 0.05$ indicates statistical significance [14].

Table 11. Analysis between MobileNetV2 and VGG16 model.

Model	Mean Accuracy (%)	SD	95% CI	p-value vs. baseline
MobileNetV2	96.8	0.7	[96.1, 97.5]	0.012 (<i>significant</i>)
VGG16	94.3	1.1	[93.2, 95.4]	—

These results, demonstrated in Table 11, show that MobileNetV2 achieves significantly higher accuracy with lower variability, confirming both consistency and robustness. Reporting such statistical metrics enhances transparency, aligns with best practices in reproducible research.

5.11. Differentiation from Existing Approaches

Equally important is how federated learning respects the confidentiality of local data as model training occurs on edge devices [9]. It does require, however, high communication bandwidth, simultaneous participation of all devices, and potentially compromising the system security with dangerous aggregation vulnerabilities [32]. In stark contrast, in our cloud-based approach, devices bear minimal training overhead while maintaining strict privacy as the device encrypted and XOR segmented images are never decrypted while being in the cloud, and all decryption occurs entirely on the device.

Equation (15) of the proposed MobileNetV2 model incorporates depthwise separable convolutions and inverted residuals to optimally spend computational resources and maximize accuracy.

$$Y = f(X) = \sigma(Wd * (Wp * X)) \quad (15)$$

where Wd and Wp denote depthwise and pointwise convolution kernels, respectively, and σ represents the non-linear activation [14]. This enables lightweight yet effective learning on cloud platforms.

Table 12. Comparative analysis of the proposed vs existing models.

Approach	Strengths	Limitations	Proposed Solution Advantage
Multimodal Authentication [15]	High robustness, anti-spoofing	High cost, latency, user resistance	Robustness via cognitive layer
Federated Learning [3,5]	Privacy-preserving training	Bandwidth-heavy, aggregation risk	Centralized XOR + lightweight DL
Proposed Method	Lightweight, cloud-scalable, secure	Limited to face + cognitive inputs	Balanced accuracy & efficiency

Hence, as summarized in Table 12, compared with other multimodal or federated approaches, our method offers a lightweight, secure, and cloud-ready solution for real-time cognitive image authentication.

5. Conclusion

This research presents an improved framework for cognitive image authentication in the cloud which incorporates deep learning and cognitive biometrics for improved security and dependability of user authentication systems. Traditional authentication methods, which are prone to spoofing and hacking, are circumvented by the proposed system which features a blend of facial recognition and cognitive biometric features. Performance and security of the system in real time have been boosted by the combination of the MobileNetV2 model and XOR encryption used for image segmentation. Results from experiments suggest that the model based on MobileNetV2 has a better performance than the standard models, attaining an accuracy of 95.3%, and F1 score 91.11%, which indicates a high level of performance on the dataset. Moreover, the system's execution time for XOR encryption of 55ms to 75ms is more than sufficient for real time applications. This work proves that the coupled user authentication systems on mobile and cloud applications can be efficiently and securely deployed using cloud integrated cognitive image authentication systems. Future refinements on the model should aim at alleviating issues caused by poor lighting, aging, and accessories for improved robustness and scalability.

All the Declarations and Statements

Author Contributions Statement

Pranali Dahiwal – Conceptualization, Methodology, and proposed research ideas, construction of the overall framework, and supervised project execution.

Vijay Khare – Writing – Review and Editing and ensured clarity and coherence.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors have no relevant conflicts of interest to disclose.

Funding Declaration

The authors received no funding support from any agencies or firms.

Data Availability Statement

This study analyzed publicly available datasets.

Ethical Declarations

The authors declare that this study did not involve any human participants or animals.

Acknowledgements

None.

Declaration of Generative AI in Scholarly Writing

None.

Abbreviations

The following abbreviations are used in this manuscript:

IAM - Identity and Access Management
 CNN - Convolutional Neural Network
 QoS - Quality of Service
 AH - Artificial Hummingbird
 PHR - Personal Health Records
 FIBE - Fuzzy Identity Biometric Encryption
 FGSM - Fast Gradient Sign Method
 PGD - Projected Gradient Descent

APPENDIX A\B\C..., with appendix title

None.

References

- [1] P. Su, "Immersive online biometric authentication algorithm for online guiding based on face recognition and cloud-based mobile edge computing," *Distributed and Parallel Databases*, vol. 41, pp. 133–154, 2023.
- [2] M. S. Hossain, G. Muhammad and M. Al Qurishi, "Verifying Image Authenticity in Cognitive Internet of Things (CIoT)-Oriented Cyber Physical System," *Mobile Networks and Applications*, vol. 23, pp. 239-250, 2018.
- [3] A. H. Sodhro, C. Sennersten and A. Ahmad, "Towards cognitive authentication for smart healthcare applications," *Sensors*, vol. 22, p. 2101, 2022.
- [4] N. I. Mowla, I. Doh and K. Chae, "On-device AI-based cognitive detection of bio-modality spoofing in medical cyber physical system," *IEEE Access*, vol. 7, pp. 2126-2137, 2019.
- [5] Z. Yu, J. Komulainen, X. Li, and G. Zhao, "Review of Face Presentation Attack Detection Competitions," in *Handbook of Biometric Anti-Spoofing*, 3rd ed., *Advances in Computer Vision and Pattern Recognition*, Springer, Singapore, pp. 287–336, 2023.
- [6] S. Bhattacharjee, A. Mohammadi, A. Anjos, and S. Marcel, "Recent Advances in Face Presentation Attack Detection," in *Handbook of Biometric Anti-Spoofing*, 2nd ed., *Advances in Computer Vision and Pattern Recognition*, Springer, Cham, pp. 207–228, 2019.
- [7] A. Kaklauskas, A. Abraham, I. Ubarte, R. Kliukas, V. Luksaite, A. Binkyte-Veliene, I. Vetloviene and L. Kaklauskienė, "A review of AI cloud and edge sensors, methods, and applications for the recognition of emotional, affective and physiological states," *Sensors*, vol. 22, p. 7824, 2022.
- [8] M. M. Ahsan, K. D. Gupta, A. K. Nag, S. Poudyal, A. Z. Kouzani and M. A. P. Mahmud, "Applications and evaluations of bio-inspired approaches in cloud security: A review," *IEEE Access*, vol. 8, pp. 180799-180814, 2020.
- [9] P. Kairouz, H. B. McMahan, et al., "Advances and Open Problems in Federated Learning," *Foundations and Trends in Machine Learning*, vol. 14, nos. 1-2, pp. 1-210, 2021.
- [10] A. T. Alrahlawee, A. D. Duru , O. Bayat and O. N. Uçan, "Cloud authentication based face recognition technique," *AURUM Journal of Engineering Systems and Architecture*, vol. 3, pp. 79-96, 2019.
- [11] S. O. Olabanji, O. O. Olaniyi, C. S. Adigwe, O. J. Okunleye and T. O. Oladoyinbo, "AI for identity and access management (IAM) in the cloud: Exploring the potential of artificial intelligence to improve user authentication, authorization, and access control within cloud-based systems," *Asian Journal of Research in Computer Science*, vol. 17, pp. 38-56, 2024.
- [12] E. A. W. Hachim, M. T. Gaata and T. Abbas, "Voice-authentication model based on deep learning for cloud environment," *International Journal on Informatics Visualization*, vol. 7, pp. 864-870, 2023.
- [13] H. Jamil, A. Ali, M. Ammi, R. Kirichek, M. S. A. Muthanna and F. Jamil, "Machine Learning-Based Identity and Access Management for Cloud Security," in *Secure Edge and Fog Computing Enabled AI for IoT and Smart Cities*, EAI/Springer Innovations in Communication and Computing, Springer, Cham, 2024, pp. 195–207.
- [14] N. Subramanian, S. Nikkath Bushra, G. Shobana, and S. Radhika, "An optimal modified bidirectional generative adversarial network for security authentication in cloud environment," *Cybernetics and Systems*, pp. 1–33, 2024.
- [15] S. Poomalai, K. Venkatesan, S. Subbaraj and S. Radha, "Secure and privacy improved cloud user authentication in biometric multimodal fusion using blockchain-based lightweight deep instance-based DetectNet," *Network: Computation in Neural Systems*, vol. 35, pp. 300-318, 2024.
- [16] C. Venkatachalam, K. Manivannan and S. Venkatachalam, "Securing data in the cloud: The application of fuzzy identity biometric encryption for enhanced privacy and authentication," *Lecture Notes in Networks and Systems*, vol. 798, pp. 213-224, 2023.
- [17] R. Shah and S. K. Dubey, "Multi user authentication for reliable data storage in cloud computing," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 10, pp. 82-89, 2024.
- [18] K. A. Patel, S. Shukla and S. J. Patel, "A novel and provably secure mutual authentication protocol for cloud environment using elliptic curve cryptography and fuzzy verifier," *Concurrency and Computation: Practice and Experience*, vol. 36, p. e7889, 2023.

- [19] M. Gupta, L. Ahuja and A. Seth, "Security enhancement in a cloud environment using a hybrid chaotic algorithm with multifactor verification for user authentication," *International Journal of Computer Applications*, vol. 45, pp. 680–696, 2023.
- [20] H. Patwal, R. Kumar, I. Ahamad, A. Mittal, H. Singh and S. Goyal, "Facial recognition in cloud security: Research perspectives on authentication solutions," *4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pp. 1869-1874, 2024.
- [21] X. Qi, C. Wu, H. Qi, Y. Shi, K. Duan and X. Wang, "A real-time face detection method based on blink detection," *IEEE Access*, vol. 11, pp. 28180–28189, 2023.
- [22] J. Chen, J. Chen, Z. Wang, C. Liang and C. W. Lin, "Identity-aware face super-resolution for low-resolution face recognition," *IEEE Signal Processing Letters*, vol. 27, pp. 645–649, 2020.
- [23] Q. Wang, T. Wu, H. Zheng and G. Guo, "Hierarchical pyramid diverse attention networks for face recognition," *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8323–8332, 2020.
- [24] Q. Cao, L. Shen, W. Xie, O. M. Parkhi and A. Zisserman, "VGGFace2: A dataset for recognising faces across pose and age," *13th IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 67-74, 2018.
- [25] Y. Zheng, D. K. Pal and M. Savvides, "Ring loss: Convex feature normalization for face recognition," *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5089-5097, 2018.
- [26] M. Sarkhoshi and Q. Li, "Cognitive graphical password based on recognition with improved user functionality," *12th International Conference on Computer Science and Information Technology (CCSIT)*, pp. 17-24, 2022.
- [27] C. Katsini, C. Fidas, M. Belk, G. Samaras and N. Avouris, "A human-cognitive perspective of users' password choices in recognition-based graphical authentication," *International Journal of Human-Computer Interaction*, vol. 35, pp. 1800–1812, 2019.
- [28] S. Wiedenbeck, J. Waters, J. C. Birget, A. Brodskiy and N. Memon, "PassPoints: Design and longitudinal evaluation of a graphical password system," *International Journal of Human-Computer Studies*, vol. 63, pp. 102-127, 2005.
- [29] P. Dunphy, J. Nicholson and P. Olivier, "Securing passfaces for description," *Proceedings of the 4th symposium on Usable privacy and security*, pp. 24-33, 2008.
- [30] M. Faraki, X. Yu, Y.-H. Tsai, Y. Suh, and M. Chandraker, "Cross-Domain Similarity Learning for Face Recognition in Unseen Domains," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 15287–15296, 2021.
- [31] J. Guo, X. Zhu, C. Zhao, D. Cao, Z. Lei, and S. Z. Li, "Learning Meta Face Recognition in Unseen Domains," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 6162–6171, 2020.
- [32] V. Shejwalkar and A. Houmansadr, "Manipulating the Byzantine: Optimizing Model Poisoning Attacks and Defenses for Federated Learning," in *Proc. Network and Distributed System Security Symp. (NDSS)*, 2021.

Authors' Profiles



Pranali Dahiwal received the B.E. and M.E. degrees in Computer Engineering. She has completed the Ph.D. degree from Savitribai Phule Pune University (S. P. Pune University). Her research interests include network security and cloud computing.



Vijay Khare is a senior professor at Savitribai Phule Pune University (SPPU) where he heads the Department of Defence and Strategic Studies and also serves as Director of the International Centre and other academic units, contributing to global engagement and interdisciplinary programs.

How to cite this paper:

Pranali Dahiwal, Vijay Khare, "Cloud-Based Cognitive Image Authentication Using Deep Learning Techniques for Secure Access", *International Journal of Computer Network and Information Security(IJCNIS)*, Vol.18, No.4, pp. 233-250, 2026. DOI:10.5815/ijcnis.2026.04.12