Journal of Al-Qadisiyah for Computer Science and Mathematics Vol.17.(1) 2025,pp.Comp 214–229

Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Secure Cloud Storage Using Multi-Modal Biometric Cryptosystem: A Deep Learning-Based Key Binding Approach

Ali Amer Abd-Aljabbar[a,*], Dalal Abdulmohsin Hammood[a], and Leith Hamid Abed[b]

[a]*Electrical Engineering Technical College, Middle Technical University, Baghdad, Iraq. Email: bbc4023@mtu.edu.iq, dalal.hammood@mtu.edu.iq*

[b]*Technical Institute of Anbar, Department of Computer Systems, Middle Technical University, Baghdad, Iraq. Email: laithhamed@mtu.edu.iq*

ARTICLEINFO

Article history:
Received: 18 /2/2025
Rrevised form: 27 /2/2025
Accepted : 6 /3/2025
Available online: 30 /3/2025

Keywords: Biometric Cryptosystem, Cloud Security, Key Binding, YOLOv8, DeepFace-VGG, Fuzzy Extractors.

ABSTRACT

Cloud computing has transformed data storage but presents security challenges, especially in authentication. Traditional passwords are vulnerable to attacks, while biometric authentication offers an alternative using fingerprints and facial recognition. However, biometric templates cannot be revoked if compromised. To address this, biometric cryptosystems integrate authentication with cryptography, though existing methods face computational and security challenges. This paper proposes a secure biometric cryptosystem for cloud storage using deep learning for biometric recognition and key binding techniques like fuzzy extractors and SHA-256 hashing. The system follows three phases: enrollment (feature extraction via YOLOv8 and DeepFace-VGG, fingerprint hashing with SHA-256, and key binding using fuzzy extractors), verification (cosine similarity for biometric matching), and encryption (AES-256 for secure storage). Experimental results show high authentication accuracy (mAP of 0.984 at 50% IoU), with FAR of 0%, FRR of 0.93%, and GAR of 99.07%. The comparative analysis highlights DeepFace-VGG's effectiveness in feature extraction, SHA-256's secure key generation, and fuzzy extractors' reliability in key reconstruction. The system resists replay attacks and biometric fluctuations while maintaining usability. Findings confirm that biometric authentication with cryptographic key binding enhances cloud security. Future research should explore scalability, multimodal biometrics, and advanced security methods like blockchain and homomorphic encryption.

MSC.

https://doi.org/10.29304/jqcsm.2025.17.11976

1.Introduction

Cloud computing is a commonly adopted computing model. It offers many advantages to users and providers. One of the most compelling reasons for its adoption is that it can be used to provide infrastructure and functionality at a reduced cost. The services are multifaceted, pay-as-you-go, on-demand, and highly effective in managing and storing data. These advantages have drawn many companies and individuals [1]. Nonetheless, the rapid expansion of cloud computing has presented considerable difficulties in safeguarding sensitive information within decentralized settings. While cloud infrastructures offer remarkable scalability and efficiency, they are susceptible to security risks stemming from their inherently shared and multi-tenant characteristics.

*Corresponding author Ali Amer Abd-Aljabbar

Email addresses: bbc4023@mtu.edu.iq

Communicated by 'sub etitor'

Traditional password-based systems have become less effective at preventing unauthorized access and reducing data breaches; therefore, alternatives like biometric cryptosystems become very appealing. Unconstrained biometrics—those that go beyond the traditional fingerprint and facial recognition, including voice, gait, and behavioral biometrics—come of particular relevance to address such challenges. These approaches leverage unique and hard-to-spoof characteristics and offer improved usability and security [2] [3].

Traditional authentication techniques- token-based systems—for example, smart cards and one-time passwords—and knowledge-based systems—including passwords and PINs—are simple in design but have serious drawbacks. Token-based systems are prone to loss, theft, or duplication, while knowledge-based systems face brute force attacks, phishing attacks, and social engineering techniques [4][5]. To overcome these weaknesses, multi-factor authentication based on biometric methods has been proposed, but those systems face challenges relating to data privacy, processing delays, and the requirement for accurate recognition in cloud environments [6]. This study enhances the cloud security field by introducing a new biometric cryptographic authentication system that leverages deep learning and secure key binding technologies. The key contributions are:

- YOLOv8-Based Face Recognition Integration: Advanced deep learning models allow for highly accurate face detection and feature identification, improving the consistency of authentication.
- Biometric Key Binding using Fuzzy Extractors: The employment of fuzzy extractors in secure auxiliary data generation allows for biometric tolerance while preventing attacks via template inversion.
- AES-256 Secured Key Management: The security is strengthened using strong encryption mechanisms that lock biometric-based cryptographic keys to protect from unauthorized use.
- Cosine Similarity-Based Verification Efficiency: The use of cosine similarity in biometric template comparison results in high matching precision and system resilience.
- Experimental Verification and Benchmarking: An in-depth performance analysis compared the proposed system to existing biometric cryptosystems to determine its usability in real-world cloud security applications.

2. Related Works

Biometric cryptosystems are the process of security verification that confirms an individual's identity through unique biological traits. During this verification process, biometric data is matched against stored datasets, therefore proving to be a strong alternative to traditional password-based systems. Typical biometric features include iris patterns, palm prints, retina scans, fingerprints, facial features, and voice signatures [7]. Recent deep-learning innovations have revolutionized the field of biometric cryptosystems with tremendous success in tasks related to machine vision, audio recognition, and natural language processing. Deep learning models, such as VGGFace, have been crucial for face recognition by providing leading accuracy in extracting hierarchical facial features in a series of tasks, from phone unlock systems to airport security checks. [8], [9]. YOLO (You Only Look Once) has benchmarked face detection for real-time performance and is proficient at identifying faces, even in challenging environments [10]. For voice-based biometrics, Long Short-Term Memory (LSTM) networks are better at capturing temporal dependencies in voice patterns to accurately verify users despite variations in speech and environmental noise [11], [12]. The above models address the emerging complexities in biometric cryptosystems through improvements in robustness against adversarial inputs, scalability for large datasets, and efficiency for real-time applications. Hence, machine learning approaches have been the building blocks in developing secure, scalable, and reliable biometric systems for various applications [13].

Abdellatef et al. [14] proposed a methodology that features the extraction of deep features from facial regions, a fusion network combining these features, and a bio-convergent approach for the maintenance of privacy and security of biometric templates. Bio-convolving encryption does not reduce the accuracy of the system; hence, it protects data reliably and in detail. The suggested method boosted accuracy, specificity, precision, recall, and F-score with biometric data protection. It also exceeded CoCo loss at 98.73% in the recognition accuracy of 98.89%, enhancing unimodal systems due to its large area under the curve. Finally, it pointed out one limitation: the performance of the facial recognition system may vary according to resolution, pose, and illumination.

Sudhakar et al. [15] applied multiple biometric methods and cross-convolution techniques to enhance security and privacy. Deep learning methods such as CNN and MLP were applied for feature extraction and user authentication.

The revocable biometrics system on the cloud was integrated with deep learning to enhance security and effectiveness. By leveraging the parallel processing of convolutional neural networks (CNNs), the system achieved a remarkable accuracy of 99.55% in the user verification task. (CB) techniques enhance user privacy and security by generating undecipherable templates. Revocable biometric data systems face challenges such as complexity, reversibility, high computational intensity, storage overhead, template deprecation, dependence on cloud infrastructure, low compatibility, data breach risk, and scalability issues.

A hybrid approach to Mobile Cloud Computing (MCC) authentication has been proposed by Zeroual et al. [16] The concepts of encryption and deep learning models have been combined. This methodology used homomorphic encryption, deep neural networks, and local ternary patterns. It has raised the accuracy of face recognition by 100% in the ORL dataset and 93.93% in the Yale dataset, with the primary concerns being data privacy and mobile resource optimization. It improved face recognition accuracy through partially homomorphic encryption, which is more efficient for mobile devices than CLBP and CLTP. Resource constraints, key complexities, network dependence, algorithmic compatibility, scalability issues, security concerns, and the availability of ample training data are major challenges the proposed hybrid solution faces.

Shukla et al. [17] Proposed a system for cancelable biometrics in which a one-way function creates the cancellable template by extracting and transforming biometric features to verify the user's identity independently. Implementing multi-factor authentication in a client-server architecture on the cloud enhanced its credibility and security, making it a robust solution for biometric-based security in the new digital landscape. The proposed cloud-based cancelable biometric systems enhanced time efficiency with faster learning of neural network models and offered secure, cancelable, noninvertible, and unlikable templates against different biometric threats. Before the widespread adoption of biometric systems, numerous issues such as hacking, data security, non-invertibility, performance, cost, environmental factors, user acceptance, and regulatory compliance must be addressed.

3. Methodology

The proposed methodology integrates biometric information with cryptographic protocols in a manner that secures information stored in a cloud environment. By utilizing biometric factors such as face and fingerprints as key security factors, the mechanism utilizes sophisticated methodologies such as Deep Learning, Binding of keys, and the AES-256 cipher algorithm in an endeavor to boost security, as shown in Figure 1. There are three critical phases in the framework: **enrolment, verification,** and **encryption**. All phases have been designed with meticulous care in an endeavor to safeguard information stored in the cloud and effectively counteract unauthorized access. The enrolment stage involves taking and registering biometric information of face and fingerprints in the system.

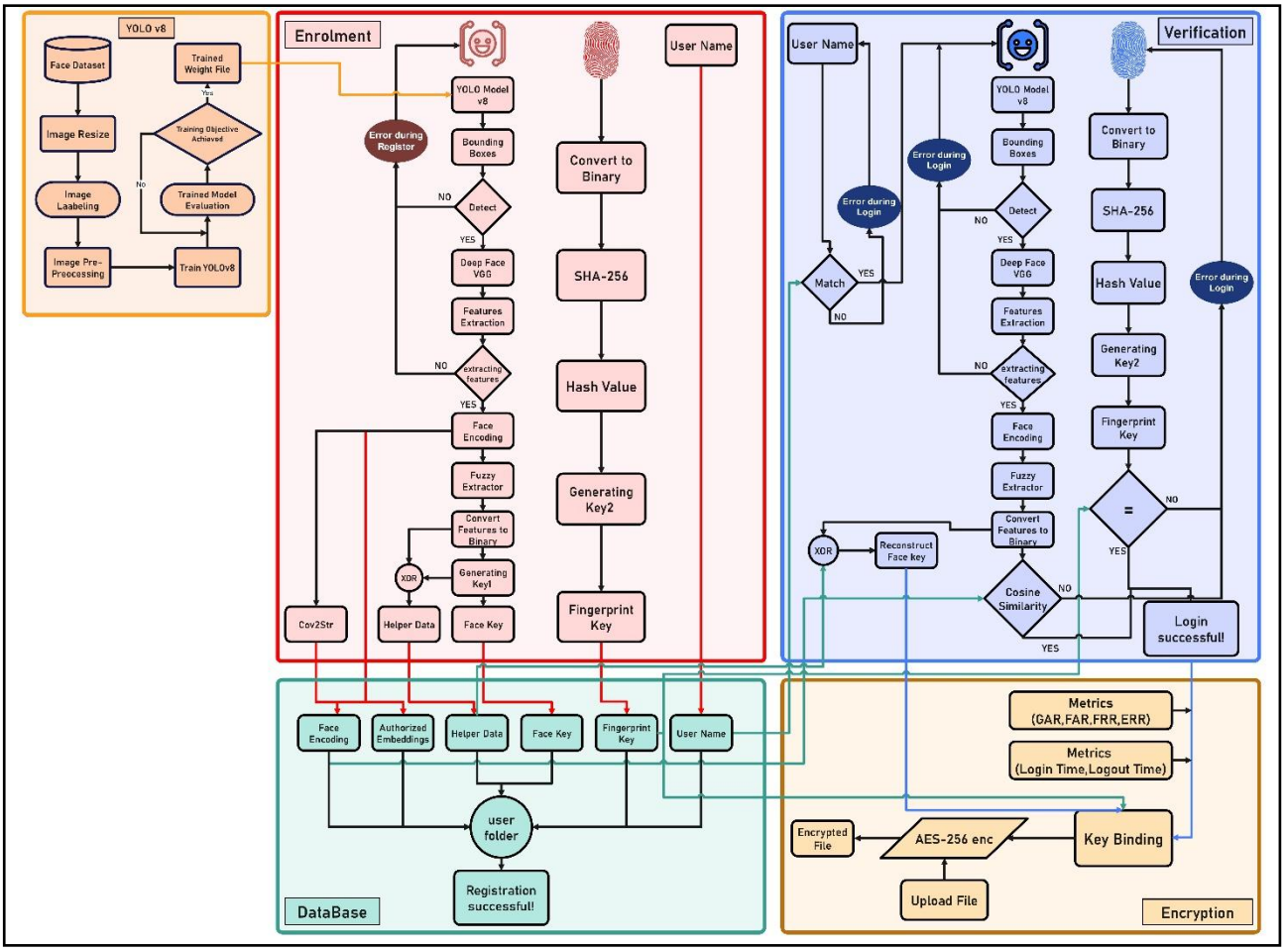


Fig. 1- General Design of the Proposed System.

3.1. Dataset and Preprocessing

This study employs face and fingerprint images for biometric authentication and encryption. The Labeled Faces in the Wild (LFW) dataset is considered a benchmark in face recognition research, consisting of 13,233 images of 5,749 subjects collected from various real-world online sources. LFW provides a diverse range of facial variations, including differences in illumination, facial angles, expressions, and occlusions, making it an essential dataset for evaluating biometric recognition systems under unconstrained conditions. Unlike controlled biometric datasets, LFW images are captured in a non-intrusive manner, meaning that subjects are not directly interacting with the capturing device. This characteristic closely aligns with the methodology of this research, where biometric data is collected passively without requiring user cooperation, introducing significant variability and noise into the dataset [87]. Acquiring non-controlled biometric data poses unique challenges to face recognition because it includes higher noise, occlusion, and blurring of motions that do not normally manifest in conventional, user-controlled systems of acquiring biometric data. This variability degrades the recognition models and, therefore, the need for the application of deep learning-based feature extraction. The combination of YOLOv8 and DeepFace-VGG in this study eased the problem of non-controlled acquisition of biometric data. The application of the strong object detection of the use of YOLOv8 aided the detection of the faces regardless of the backgrounds they had appeared in, and the application of DeepFace-VGG aided the extraction of stable and variation-insensitive feature embeddings against occlusion, illumination, and pose variation. Also, Cosine Similarity has been used as the primary matching algorithm in an attempt to improve the recognition rates with the extra intra-class variation that comes with the passive acquisition of the biometric information. This aligns with the general scope of the study, which is developing a non-intrusive yet highly secure system of authentication that will be able to perform effectively within the real world and in non-controlled environments.

To conduct the system evaluation with experiments, a subset of 17 persons of the LFW database with 100 images per subject was utilized. This subset allowed the chance for the comprehensive testing of the system's accuracy and

robustness against real-world variation of the biometric and the effectiveness of the proposed technique regardless of the non-intrusive acquisition of the biometric.

To enhance detection accuracy, the YOLOv8 model is utilized for face detection, enabling precise cropping and preprocessing of facial regions for further feature extraction using DeepFace-VGG. This integration ensures reliable biometric verification and enhances overall system accuracy. The dataset is systematically split into three subsets: 70% for training, 20% for validation, and 10% for testing. This partitioning ensures effective model generalization and prevents overfitting. Moreover, a balanced distribution across 17 classes is maintained to avoid bias and improve classification consistency. Preprocessing techniques include resizing images to 640x640 pixels to standardize input dimensions. Data augmentation techniques such as horizontal flipping, brightness adjustments, exposure modifications, and noise injection are applied to improve model robustness against variations in image quality and environmental conditions. For accurate labeling and annotation, the Roboflow [19] The platform is used to incorporate bounding boxes and automated label assistance to refine object localization. This step ensures high-quality input data for training deep learning models.

For fingerprint authentication, the FVC2000_DB4 database is the largest and the most widely used database for commercial and scholarly research in the evaluation of fingerprint recognition system performance. The database has been released in the scope of the Fingerprint Verification Competition 2000, and it comprises 800 fingerprint images that belong to 100 different individual fingerprints with a total of eight samples per fingerprint [89].

One of the characteristic features of this database is the heavy noise and distortion in the images, which thus form a challenging benchmarking database for fingerprint recognition systems. The image database includes a mixture of artifacts such as sensor noise, distortion, and blurring that severely negatively influence the matching and accuracies of the features. This has been caused primarily by the non-ideal acquisition environment and the use of a thermal sweeping sensor in the fingerprint acquisition. The resolution of the sensor is 500 dpi, and it also has contrast variation, ridge discontinuity, and partial print that further raises the recognition task difficulty.

SHA-256 hashing during fingerprint template generation ensures good security with the maintenance of matching accuracy regardless of such variances. Cosine Similarity also ensures a good measure of addressing variances of the biometric templates and diminishes the influence of noise and distortion.

This database remains a valuable source of images against which fingerprint and minutiae-based matching algorithmic testing and testing of identity verification and biometric encryption systems may be performed. Due to the challenging variation of image quality and the small surface area of the scans within the FVC2000_DB4 database, the database forms a good benchmark against the strength and efficiency of multiple-biometric security systems that may be assessed based on a real-world scenario with noise and distortion.

3.2. Enrolment Phase

The enrollment step is vital for secure biometric data registration. Face detection is performed by the system through YOLOv8, followed by feature extraction through DeepFace-VGG to create distinctive face encodings. Face encodings go through a Fuzzy Extractor to convert to binary to create helper data and face key. Simultaneously, SHA-256 hashing is used on fingerprint data to develop a cryptographic fingerprint key. Helper data and biometric keys are saved securely in the database for future encryption and verification.

- DeepFace-VGG: It is a deep model that extracts facial features of higher dimensions through convolutional layers to create robust embedding that guarantees secure biometric verification. Embedding is saved or compared through Cosine Similarity in verification.

- Fuzzy Extractor: It maximizes security by transforming facial feature-extracted facial features into binary vectors. The random cryptographic key is produced and XORed with this vector to create secure helper data. Key generation on demand is enabled through this method while biometric template protection is ensured.

- SHA-256 for Fingerprint Processing: Fingerprint feature vectors get normalized before going through SHA-256 hashing to create a 256-bit fingerprint key. Securely saved, it is used for verification.

- Database Structure: MySQL database is used to store biometric data in a secure manner that includes face encodings, fingerprint hashes, cryptographic keys, and helper data. Secure retrieval of biometric

credentials is ensured through the database, while sensitive details get hashed and encrypted for the protection of confidentiality and integrity of data.

3.3. Verification Phase

The verification step ensures secure and accurate user authentication by face and fingerprint biometric authentication. To begin with, the user enters their face image and their fingerprint. For face verification, the face is detected by YOLOv8, whereas high-dimensional features are extracted by DeepFace-VGG and converted to binary format. The face key is reconstructed by a Fuzzy Extractor and matched to stored embeddings by cosine similarity. If the score of similarity is lower than set beforehand, access is denied. For fingerprint matching, the presented fingerprint is converted to a binary format, whereas SHA-256 hashing results in a cryptographic key that is compared to the stored fingerprint hash in the database. Access is denied in the event of a mismatch. Authentication is only granted if face and fingerprint verifications succeed in offering a highly secure biometric login technique. The given approach integrates YOLOv8, DeepFace-VGG, SHA-256, and Fuzzy Extractors to offer solid reliability and robust protection. Cosine Similarity in Face Verification Cosine similarity is employed to determine how similar two vectors are by computing the cosine of their separation angle [21]. During login, DeepFace-VGG loads a facial feature vector that is compared to stored embeddings by utilizing the following formula:

- Dot Product Calculation:

$$A \cdot B = \sum_{i=1}^{n} A_i \cdot B_i \tag{1}$$

- Magnitude of Vectors:

$$|A| = \sqrt{\sum_{i=1}^{n} A_i^2} \tag{2}$$

$$|B| = \sqrt{\sum_{i=1}^{n} B_i^2} \tag{3}$$

- Cosine Similarity Formula:

$$\boldsymbol{C}osine\ Similarity = \frac{A \cdot B}{\| A \| \| B \|} \tag{4}$$

A similarity score near 1 indicates a strong match, while a lower score results in login rejection. A predefined threshold of 0.8 determines whether authentication is successful. If the score is below this threshold, access is denied. By combining face biometric technology with fingerprint technology, the system has secure biometric protection that is of high reliability and anti-tamper in nature.

3.4. Encryption and Decryption

The techniques for decryption and encryption discussed in the proposed scheme serve a critical function in protecting the integrity and confidentiality of stored and communicated user information in cloud environments. Biometric key generation, in combination with strong cryptographic algorithms, strengthens information security through protection against unauthorized access. As shown in figure 2.

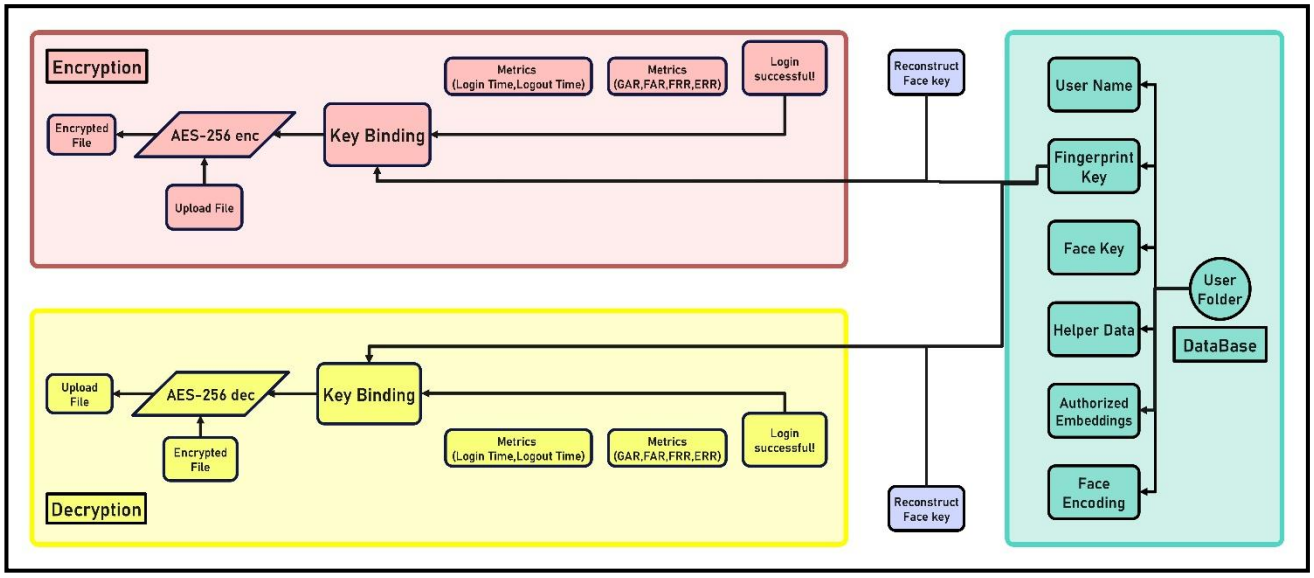


Fig. 2- Encryption and Decryption scheme in proposed methodology.

According to the envisaged framework, the key binding process forms a critical mechanism for combining keys derived from biometric sources, namely faces and fingerprints, into a single key for use in both encryption and decryption processes. It grants effective security for sensitive information stored in cloud storage infrastructure. In addition, it maximizes the uniqueness and accuracy of biometric information and incorporates cryptographic techniques for countering any potential unauthorized access vulnerabilities.

AES 256 (Advanced Encryption Standard) is one of the most widely used block encryption algorithms today since it has the strength and efficiency to bear all kinds of attacks. This algorithm works based on dividing plaintext into 128-bit blocks with the aid of a 256-bit encryption key. The entire encryption process consists of 14 iterations, which make use of four basic operations: SubBytes, ShiftRows, MixColumns, and AddRoundKey. Each subsequent pass adds to the complexity of the encrypted message, making it nearly impossible to deduce the original text without possessing the decryption key. [22].

- SubBytes: Bytes are substituted according to a non-linear transformation table $S(x)$ to provide security against differential analysis [23]:

$$S'(i,j) = S\big(S(i,j)\big) \tag{5}$$

$S'(i,j)$: The new value of the byte at row i and column j after applying the substitution box (S_{box}). $S(i,j)$: The original value of the byte at row i and column j. S: The Substitution Box is a non-linear transformation designed to resist differential and linear cryptanalysis.

- ShiftRows: Based on a shift in which rows in the state grid. The first row is an unchanged second row is shifted by one byte, the third by two, and the fourth by three bytes [23]:

$$S'(i,j) = S\big(i,(j + shift(i))mod\ 4\big) \tag{6}$$

S′(i, j): The new value of the byte at the position (i, j) after row shifting. S(i, j): The original value of the byte at the position (i, j). shift(i): The number of positions to shift row i:
- Row0(i = 0): No shift.
- Row1(i = 1): Shift by 1 position.
- Row2(i = 2): Shift by 2 positions.
- Row3(i = 3): Shift by 3 positions.

mod 4: Ensures the shifting stays within the four columns.

- MixColumns: A linear mixing operation wherein the columns of the state grid are multiplied by a constant matrix M via the operations of multiplication and modulo. It's performed on finite fields GF(2^8) [23]:

$$S'(i,j) = M \cdot S(i,j) mod x^8 + x^4 + x^3 + x + 1 \quad (7)$$

$S'(i,j)$: The new value of the byte at the position (i,j) after column mixing. $S(i,j)$: The original value of the byte at the position (i,j). M: A fixed matrix $(4x4)$ used for linear mixing through multiplication in the finite field $GF(2^8)$. Typically, the matrix is:
M =

$$\begin{bmatrix} 2\,3\,1\,1 \\ 1\,2\,3\,1 \\ 1\,1\,2\,3 \\ 3\,1\,1\,2 \end{bmatrix}$$

- AddRoundKey: XOR operation is applied between the current block and the round key K_{round} [23]:

$$S'(i,j) = S(i,j) \oplus K_{round(i,j)} \quad (8)$$

$S'(i,j)$: The new value of the byte at the position (i,j) after applying the round key. $S(i,j)$: The original value of the byte at the position (i,j). $K_{round}(i,j)$: The value of the byte at position (i,j)In the current round key. $\oplus$: The XOR (Exclusive OR) operation.

- Key Schedule: The keys used in the rounds are generated from the base encryption key via a key expansion function. The algorithm uses 8 S_{Boxes} per round to generate the round key [23]:

$$k_{\{l+1,i,j\}} = k_{\{l,i,j-1\}} \oplus k_{\{l,i,j\}} \quad (9)$$

$k_{\{l+1,i,j\}}$: The value of the byte at position (i, j) in the key for round $l + 1$. $k_{\{l,i,j-1\}}$: The value of the byte at position $(i, j - 1)$ ***In*** the key for round l. $k_{\{l,i,j\}}$: The value of the byte at position (i, j) ***In*** the key for round l. $\oplus$: The XOR operation.

3.5. Evaluation Metrics

In order to assess the effectiveness and reliability of the biometric-based encryption and authentication scheme, a range of evaluation metrics is adopted. These allow for the accuracy, robustness, and security of the scheme to be measured and, in turn, validate its usability in actual implementations. These metrics allow us to assess the models' performance in the following ways:

- **IoU**: Use IoU to calculate how closely the predicted and actual bounding boxes coincide. It considers two bounding boxes' overlap with regard to their overall area. To verify accuracy, contrast a predicted bounding box with the correct bounding box and view its IoU value. To calculate the IoU value, divide both bounding boxes' common area with regard to both bounding boxes' overall area.

- **Precision**: Precision [24] Functions as a key performance metric for proposed biometric security and authentication systems, specifically its ability to accurately discriminate between actual persons with a minimum of incorrect acceptances (False Positives, FP). Precision is a ratio of accurately detected actual positive cases, including both real and imposter cases, to the total number of positively detected cases. High values of precision represent high performance in distinguishing between actual persons and unauthorized access and, in consequence, a reduced chance of incorrect acceptances (False Positives, FP). Precision can be calculated using Equation 10:

$$Precision = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \quad (10)$$

- **Recall:** Recall [25] Evaluates a biometric system's ability to effectively authenticate a valid user and, at the same time, minimize cases of incorrect rejections. Recall measurement is a function of the proportion of True Positives (TP) to actual positive cases, including False Negatives (FN). A high value for Recall reflects a low rejection level for valid users, an important feature in scenarios with biometric encryption and key binding. Nevertheless, recall and precision must be balanced in terms of security concerns. Recall calculation is represented in Equation 11:

$$Recall = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)} \quad (11)$$

- **Mean Average Precision (mAP):** Mean Average Precision (mAP) [26] It is a critical performance evaluation for models in object detection, offering an effectiveness estimation through a consideration of precision and recall at a range of Intersections over Union (IoU) thresholds. It begins with a calculation of Average Precision (AP) for one individual category, with AP estimating the area under a precision-recall curve. mAP is then calculated through aggregation of AP values for all defined object classes. High mAP reflects a high level of model effectiveness, with an efficient system in its performance at identifying objects with a minimum of both false positives and false negatives. mAP is regularly used in deep architectures, such as YOLO, Faster R-CNN, and SSD, for evaluation. mAP calculation is represented in Equation 12:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (12)$$

In this regard, AP_i stands for average precision with regard to the i-th category, and N is a variable representing the number of object classes in its totality. Quantitatively representing Intersection over Union (IoU) in terms of mean Average Precision (mAP) is a widespread practice. For example, mAP50−95 describes a range of IoU values between 0.5 and 0.95 with a 0.05 interval, whereas mAP95 describes a single value for IoU, 0.95. As seen in Equation 10, the computation of average precision for a target class begins with ranking all detected items in a descending manner according to confidence values. Next, at each level, recall and precision values are computed, culminating in the computation of integration of the precision-recall curve over all accessible thresholds in an estimation of an area under a curve, specifically the precision-recall curve (AUC).

$$AP_i = \int_0^1 P(R)\ dR \quad (13)$$

- **Genuine Acceptance Rate (GAR):** GAR is the percentage of cases where the system correctly identified an authorized user, as seen in Equation 14. It measures how much the system can be relied on to correctly identify authorized users [27].

$$GAR = \left(\frac{Number\ of\ genuine\ matches}{Total\ number\ of\ genuine\ attempts}\right) * 100 \quad (14)$$

- **False Acceptance Rate (FAR):** FAR is the percentage of instances where an unauthorized user is incorrectly accepted by the system, as seen in Equation 15. It is a serious security error, as it measures how the system is exposed to giving access to impostors [28].

$$FAR = \left(\frac{Number\ of\ false\ matches}{Total\ number\ of\ impostor\ attempts}\right) * 100 \quad (15)$$

- **False Rejection Rate (FRR):** FRR Percentage of instances where an authorized user is falsely rejected by the system, as seen in Equation 16. This represents a functional error where legitimate users are denied access [29].

$$FRR = \left(\frac{Number\ of\ false\ rejections}{Total\ number\ of\ genuine\ attempts}\right) * 100 \quad (16)$$

- **The Equal Error Rate (EER):** is an important metric in biometric authentication, representing the point at which both the False Rejection Rate (FRR) and False Acceptance Rate (FAR) become equal [30]. A smaller EER indicates a higher level of accuracy in the system, attaining an ideal balance between security and usability. It is typically derived from the Receiver Operating Characteristic (ROC) curve. The formula for EER is:

$$EER = FAR(\tau) = FRR(\tau) \quad (17)$$

Where τ is the decision threshold at which both rates are equal. EER is widely used to compare biometric systems, with lower values indicating improved performance.

4. Results and Discussion

The study was performed in two computing settings: Google Colaboratory and a computer installed on-site, with each having unique benefits for model training in deep learning and testing.

Collaboratory provides unlimited Tesla T4 GPU at no cost to significantly improve performance in the training of deep learning models. Testing was performed on the environment of Python 3.11.11, PyTorch 2.5.1, and CUDA 12.4 for performance maximization. The face detection model and feature extraction model for YOLOv8 were deployed through the Ultralytics framework. Tesla T4 offers high computing power that maximizes neural network training and inference.

Table 1 - Google Colab Experimental Environment

Experiment platform	Python	Ultralytics	Pytorch	CUDA	GPU
Google Colaboratory	Version 3.11.11	Version 8.1.5	Version 2.5.1	Version 12.4	Tesla T4

A high-performance local computer was used for further testing with controlled conditions of processing. Having Windows 10 Pro operating system, it comes with an Intel Core i7-10700K processor (8 cores, 16 threads, 3.8GHz base, turbo boost of 5.1GHz), 32GB of DDR4 at 3200MHz of RAM, and an NVIDIA RTX 3060ti GPU with 4864 CUDA cores and 8GB of GDDR6 VRAM. All these specifications make for good parallel processing of deep neural networks for their training as well as inference.

Table 2 - Local Experimental Environment.

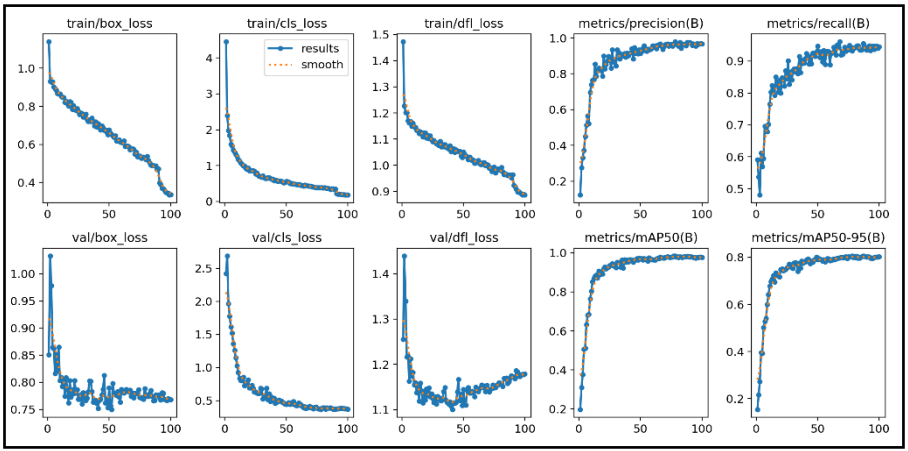
Experiment platform	Python	Windows	RAM	CPU	GPU
Local PC	Version 3.12.7	Version 10 Pro	32 GB	Intel i7-10700k	RTX 3060ti

Both environments offset one another, with cloud-based flexibility from Google Colab and on-premise infrastructure to offer performance-tuned computing for intensive computations.

4.1. YOLOv8 Results and Analysis

YOLOv8 performance is shown by train and valid loss functions along with key performance metrics. The train box loss started well over 1.0 but successively reduced to 0.3309, suggesting better bounding box localization. The train class loss also reduced from 4.0 to 0.1837, suggesting better classification, whereas train dynamic focal loss reduced from 1.5 to 0.8777, suggesting better management of challenging cases. For validation, the box loss reached 0.75, whereas the class loss and dynamic focal loss reached 0.5 and 1.1, respectively, suggesting that the model is well generalizing to unseen instances. The trends here suggest stable training without overfitting, maximizing feature extraction of faces by YOLOv8 efficiently.

The final three subplots reflect precision, recall, and mean average precision (mAP), providing insight into how accurately the model is operating. The precision of 0.981 means that 98.1% of identified objects were accurately identified, maintaining false positives to minimum levels. The 0.932 recall means that 93.2% of real objects were actually picked up, maintaining false negatives to minimum levels. mAP@50 also hit 0.984, maintaining good performance on object detection using 50% Intersection over Union (IoU) thresholds, while mAP@50-95 hit 0.807,



indicating steadiness on the fluctuating difficulty of detection. All of these ensure that YOLOv8 is effective in biometric use cases, most notably face detection and identification, maintaining maximum reliability in actual use cases. As shown in figure 3.

Fig. 3- YOLOv8 Training and Validation Performance Metrics.

The confusion matrix outlines the relationship between the predictions made by the YOLOv8 model (shown on the vertical axis) and the true values (shown on the horizontal axis) for different categories. The matrix highlights the model's strong ability in classification, as evident from the high concentration of correct predictions along the diagonal. Notably, the model showed exceptional accuracy for Scarlett Johansson, with 31 correct predictions, and for Hugh Jackman and Jennifer Lawrence, with both achieving 22 correct predictions. However, there were slight cases of misclassification in categories like Brad Pitt and Denzel Washington, where few samples were wrongly classified as belonging to other classes. Such errors often occurred in visually similar categories or with low-quality images. Overall, the matrix shows the model's high classification accuracy while identifying areas of potential improvement by addressing specific areas of confusion, enhancing data quality, and increasing the sample sizes in categories that showed suboptimal performance, as shown in Figure 4.

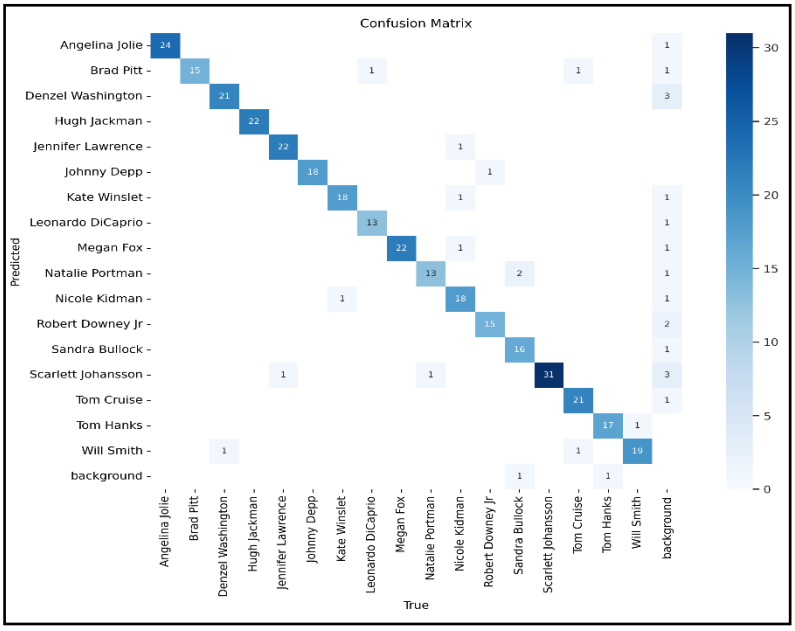


Fig. 4- Confusion Matrix for YOLOv8 Model.

4.2. Secure Key and Time Execution

A cryptographic system based on biometric information requires a delicate balance between security, usability, and reliability. In order to maintain this balance, a number of key performance metrics are measured, namely the Genuine Acceptance Rate (GAR), False Acceptance Rate (FAR), False Rejection Rate (FRR), and Equal Error Rate (EER). Figure 5 presents a complete analysis of these metrics:

Table 3 - Performance Evaluation of the Proposed Biometric Cryptosystem.

Metrics	GAR	FAR	FRR	EER
Result	99.07%	0%	0.93%	0.465%

- The GAR (99.07%) indicates that the system shows a high proficiency in correctly identifying and verifying genuine users in almost all cases.

- The FAR (0%) indicates a zero error in recognizing unauthorized users, improving the system's immunity to false positives.
- The FRR (0.93%) indicates that a small percentage of genuine users were rejected, possibly due to small inaccuracies in biometric information.
- The EER (0.465%) is a midpoint between FAR and FRR, indicating that the system maintains a balance between security and usability at its optimal point.

These metrics verify that the biometric-based cryptography system attains a high degree of precision without sacrificing security.

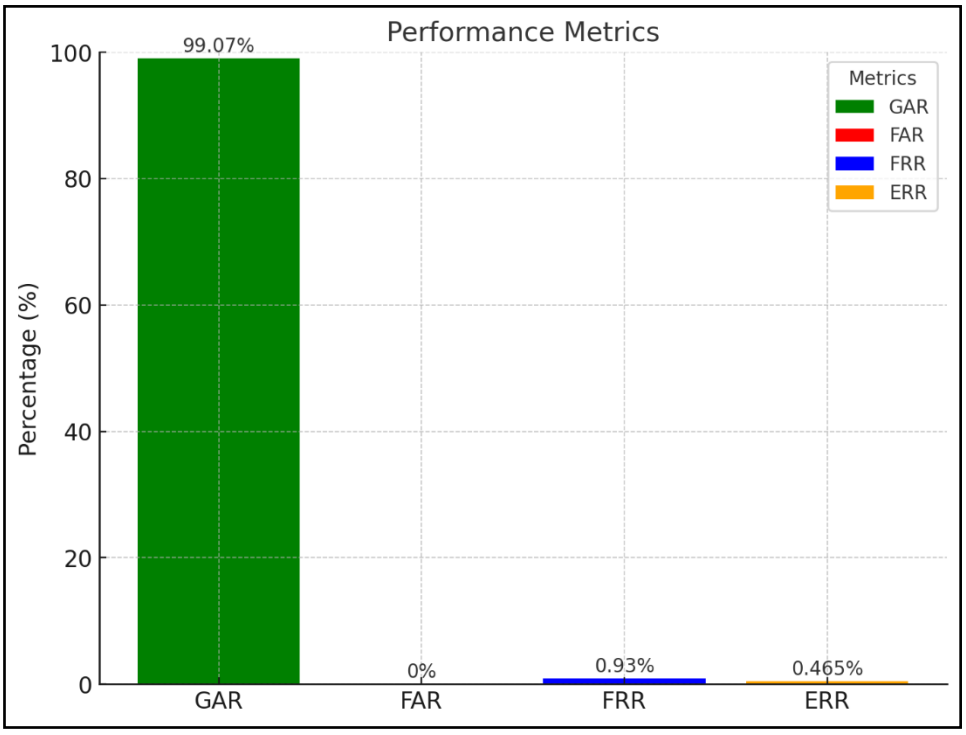


Fig. 5- Performance Metrics Analysis: GAR, FAR, FRR, and EER.

In practical applications, biometric-based systems need to work in short response times to provide an optimal user experience. The time taken to execute crucial processes is discussed in Figure 6, stressing the time taken to execute the following processes:

- Registered Embedding: The process of extracting biometric features from input data.
- Creation of a binary vector: The conversion of the extracted features to a binary vector that is usable in a cryptographic process.
- Key Generation: The conversion of a binary vector to a key that is usable in a cryptographic process for purposes of authentication.
- Creation of Helper Data: The generation of auxiliary data that is used to rebuild the key during a verification process.

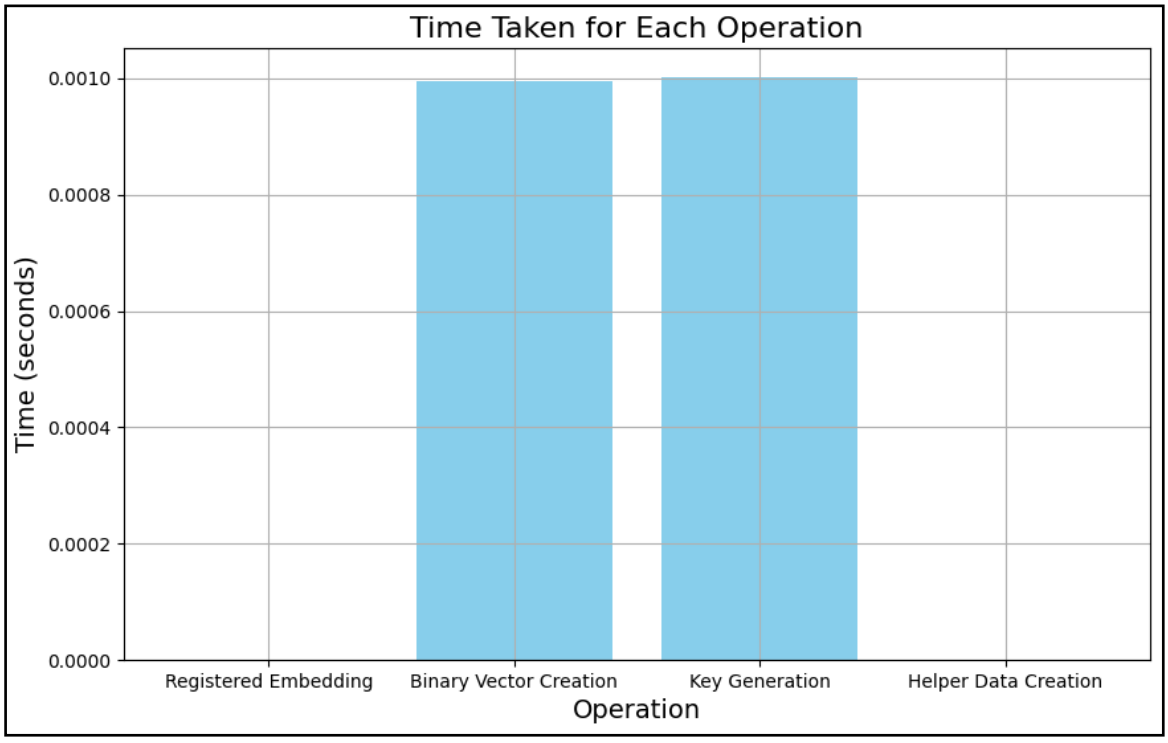


Fig. 6- Execution Time for Each Step in Biometric Cryptographic Processing.

4.3. Comparison of the proposed approach with recent studies

The proposed system integrates face detection through YOLOv8 and feature extraction through DeepFace-VGG along with fingerprint identification to ensure greater biometric authentication security. The system is trained and validated on the Labeled Faces in the Wild face recognition dataset and FVC2000 fingerprint-based authentication dataset. With feature extraction through deep learning combined with cryptographic security protocols, the system had 99.07% accuracy, indicating better performance of biometric authentication. The multi-biometric fusion of face and fingerprint, the use of deep learning-based models, and the optimization of the process justified the reliability and resilience of the system.

Table 4 - Comparing our Proposed System with Previous Studies based on Artifactual Intelligence.

Ref\ Year	Biometric Type	Feature Extraction	Dataset	Result
Abdellatef et al. [14]\ 2019	Face	CNNs	FERET	Accuracy of 98.89%
			LFW	Accuracy of 98.93%
			PaSC	Accuracy of 97.38%
Sudhakar et al. [15]\2020	Iris \ Finger veins	CNN	IITD	Accuracy of 98%
			MMU	Accuracy of 92%
Zeroual et al. [16]\2022	Physical \ behavioural	DCNN + LTP	Yale	Accuracy of 90.90%
			Georgia Tech	Accuracy of 98.66%
			FEI	Accuracy of 98.03%

			ORL	Accuracy of 98.75%
			Extended Yale	Accuracy of 98.78%
Shukla et al. [17]\2023	Iris \ Finger vein	CNNs	IITD	Accuracy of 92.1%
			MMU	Accuracy of 85.4%
			FV-USM	Accuracy of 92.3%
Our Proposed System (2025)	**Face \ Fingerprint**	**YOLOv8 + Deepface_VGG**	**LFW \ FVC2000**	**Accuracy of 99.07%**

When compared to earlier literature, our system outcompetes many of the top-notch models. Abdellatef et al. (2019), employing CNN-based face recognition models on FERET, LFW, and PaSC datasets, have reported 97.38-98.93% accuracies. Sudhakar et al. (2020), employing CNN-based iris and finger vein recognition, reported up to 98% on the IITD dataset, while Zeroual et al. (2022), employing combined DCNN and LTP on behavioral and physical biometrics, reported 98.66% on the Georgia Tech dataset. More recently, when employing iris and finger vein recognition, Shukla et al. (2023) reported 85.4-98.78% accuracies on datasets that they used.

5. Conclusion

The integration of key binding through biometrics makes cloud storage secure by overcoming weaknesses in password-based and token-based schemes. With the high-level security risks of unauthorized use and identity theft in cloud computing, biometric-based authentication combined with cryptographic key binding has been demonstrated to be an effective scheme of security. The present work introduces novel face recognition-based biometric-based authentication using face recognition (YOLOv8 and DeepFace-VGG) combined with hashing of fingerprints using SHA-256. The system is structured into three processes: registration, verification, and encryption, offering greater accuracy and attack resilience. Fuzzy extractors bind biometric features into cryptographic keys securely, protecting against replay attacks and coping with variations of biometrics through time. The experimental findings confirmed the reliability of the system, providing mean average precision (mAP) of 0.984 along with zero false acceptance rate (FAR), false rejection rate (FRR) of 0.93%, equal error rate (EER) of 0.465%, and genuine acceptance rate (GAR) of 99.07%. The integration of the AES-256 encryption scheme also provided protection of data. The comparative study confirmed that the system is very secure, efficient, and resilient to variations of biometrics, making it very effective to use in cloud-based security schemes.

Acknowledgments

We would like to express our gratitude to all the individuals and institutions who supported and contributed to this research. Some recent references were accessed through institutional subscriptions and academic collaborations.

References

[1] B. Mahalakshmi and B. David, "An Analytical Survey on Multi-Biometric Authentication System for Enhancing the Security Levels in Cloud Computing," in *Proceedings of 8th IEEE International Conference on Science, Technology, Engineering and Mathematics, ICONSTEM 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ICONSTEM56934.2023.10142265.

[2] A. M. Mostafa *et al.*, "Strengthening Cloud Security: An Innovative Multi-Factor Multi-Layer Authentication Framework for Cloud User Authentication," *Applied Sciences 2023, Vol. 13, Page 10871*, vol. 13, no. 19, p. 10871, Sep. 2023, doi: 10.3390/APP131910871.

[3] A. K. Jain, A. A. Ross, and K. Nandakumar, "Introduction to Biometrics," 2011, doi: 10.1007/978-0-387-77326-1.

[4] O. Umoren, R. Singh, Z. Pervez, and K. Dahal, "Securing Fog Computing with a Decentralised User Authentication Approach Based on Blockchain," *Sensors 2022, Vol. 22, Page 3956*, vol. 22, no. 10, p. 3956, May 2022, doi: 10.3390/S22103956.

[5] T. Grance and W. Jansen, "Guidelines on Security and Privacy in Public Cloud Computing," 2011, doi: 10.6028/NIST.SP.800-144.

[6] N. K. Ratha and V. Govindaraju, "Advances in biometrics: Sensors, algorithms, and systems," *Advances in Biometrics: Sensors, Algorithms and Systems*, pp. 1–503, 2008, doi: 10.1007/978-1-84628-921-7/COVER.

[7] A. S. Amsalam, A. Al-Naji, A. Y. Daeef, and J. Chahl, "Computer Vision System for Facial Palsy Detection," *Journal of Techniques*, vol. 5, no. 1, pp. 44–51, Mar. 2023, doi: 10.51173/JT.V5I1.1133.

[8] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 07-12-June-2015, pp. 815–823, Oct. 2015, doi: 10.1109/CVPR.2015.7298682.

[9] S. A. Salih, S. K. Gharghan, J. F. Mahdi, and I. J. Kadhim, "Lung Diseases Diagnosis-Based Deep Learning Methods: A Review," *Journal of Techniques*, vol. 5, no. 3, pp. 158–173, Sep. 2023, doi: 10.51173/JT.V5I3.1469.

[10] N. N. Ali, A. Hameed, A. G. Perera, and A. Al_Naji, "Custom YOLO Object Detection Model for COVID-19 Diagnosis," *Journal of Techniques*, vol. 5, no. 3, pp. 92–100, Sep. 2023, doi: 10.51173/JT.V5I3.1174.

[11] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, "Speech Recognition Using Deep Neural Networks: A Systematic Review," *IEEE Access*, vol. 7, pp. 19143–19165, 2019, doi: 10.1109/ACCESS.2019.2896880.

[12] A. Graves, A. R. Mohamed, and G. Hinton, "Speech Recognition with Deep Recurrent Neural Networks," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 6645–6649, Mar. 2013, doi: 10.1109/ICASSP.2013.6638947.

[13] N. Siroya and M. Mandot, "Biometric identification using deep learning for advance cloud security," in *Deep Learning Approaches to Cloud Security*, wiley, 2021, pp. 1–14. doi: 10.1002/9781119760542.ch1.

[14] E. Abdellatef, N. A. Ismail, S. E. S. E. Abd Elrahman, K. N. Ismail, M. Rihan, and F. E. Abd El-Samie, "Cancelable multi-biometric recognition system based on deep learning," *Visual Computer*, vol. 36, no. 6, pp. 1097–1109, Jun. 2020, doi: 10.1007/s00371-019-01715-5.

[15] T. Sudhakar and M. Gavrilova, "Cancelable Biometrics Using Deep Learning as a Cloud Service," *IEEE Access*, vol. 8, pp. 112932–112943, 2020, doi: 10.1109/ACCESS.2020.3003869.

[16] A. Zeroual, M. Amroune, M. Derdour, and A. Bentahar, "Lightweight deep learning model to secure authentication in Mobile Cloud Computing," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 9, pp. 6938–6948, Oct. 2022, doi: 10.1016/j.jksuci.2021.09.016.

[17] R. Shukla and H. Kaur, "Deep learning based cancelable biometric system," in *2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ICCCNT56998.2023.10307990.

[18] E. Gonzalez-Sosa, J. Fierrez, R. Vera-Rodriguez, and F. Alonso-Fernandez, "Facial Soft Biometrics for Recognition in the Wild: Recent Works, Annotation, and COTS Evaluation," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 8, pp. 2001–2014, Oct. 2022, doi: 10.1109/TIFS.2018.2807791.

[19] S. G. E. Brucal, L. C. M. De Jesus, S. R. Peruda, L. A. Samaniego, and E. D. Yong, "Development of Tomato Leaf Disease Detection using YoloV8 Model via RoboFlow 2.0," *GCCE 2023 - 2023 IEEE 12th Global Conference on Consumer Electronics*, pp. 692–694, 2023, doi: 10.1109/GCCE59613.2023.10315251.

[20] C. Gottschlich and S. Huckemann, "Separating the real from the synthetic: minutiae histograms as fingerprints of fingerprints," *IET Biom*, vol. 3, no. 4, pp. 291–301, Dec. 2014, doi: 10.1049/IET-BMT.2013.0065.

[21] A. R. Lahitani, A. E. Permanasari, and N. A. Setiawan, "Cosine similarity to determine similarity measure: Study case in online essay assessment," *Proceedings of 2016 4th International Conference on Cyber and IT Service Management, CITSM 2016*, Sep. 2016, doi: 10.1109/CITSM.2016.7577578.

[22] A. Biryukov and D. Khovratovich, "Related-Key Cryptanalysis of the Full AES-192 and AES-256," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5912 LNCS, pp. 1–18, 2009, doi: 10.1007/978-3-642-10366-7_1.

[23] A. Biryukov, O. Dunkelman, N. Keller, D. Khovratovich, and A. Shamir, "Key recovery attacks of practical complexity on AES-256 variants with up to 10 rounds," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6110 LNCS, pp. 299–319, 2010, doi: 10.1007/978-3-642-13190-5_15.

[24] F. Abbaas and G. Serpen, "Evaluation of biometric user authentication using an ensemble classifier with face and voice recognition," May 2020, Accessed: Feb. 06, 2025. [Online]. Available: https://arxiv.org/abs/2006.00548v1

[25] M. Mageshbabu and J. Mohana, "Enhancing Biometric Security: A Machine Learning Approach to ECG-Based Authentication," *2nd International Conference on Intelligent Cyber Physical Systems and Internet of Things, ICoICI 2024 - Proceedings*, pp. 1654–1659, 2024, doi: 10.1109/ICOICI62503.2024.10696328.

[26] R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. Da Silva, "A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit," *Electronics 2021, Vol. 10, Page 279*, vol. 10, no. 3, p. 279, Jan. 2021, doi: 10.3390/ELECTRONICS10030279.

[27] K. Nandakumar and A. K. Jain, "Biometric Template Protection: Bridging the performance gap between theory and practice," *IEEE Signal Process Mag*, vol. 32, no. 5, pp. 88–100, Sep. 2015, doi: 10.1109/MSP.2015.2427849.

[28] U. Uludag, S. Pankanti, and A. K. Jain, "Fuzzy Vault for Fingerprints," *Lecture Notes in Computer Science*, vol. 3546, pp. 310–319, 2005, doi: 10.1007/11527923_32.

[29] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, "Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data," *SIAM Journal on Computing*, vol. 38, no. 1, pp. 97–139, Feb. 2006, doi: 10.1137/060651380.

[30] J. M. Cheng and H. C. Wang, "A method of estimating the equal error rate for automatic speaker verification," *2004 International Symposium on Chinese Spoken Language Processing - Proceedings*, pp. 285–288, 2004, doi: 10.1109/CHINSL.2004.1409642.