Federated Multi-Modal Biometrics: A Privacy-Centric Approach to Face and Eye-Blink Based Authentication

MUTAKANI PETER[1], SHAIK LAL JOHN BASHA[2], PALEPOGU.RAJASEKHAR[3],CHENNEBOYINA CHANDRAKALA[4], Dr.K.PRAVEENA[5]

Electronics & communication engineering, Chalapathi Institute of Engineering & Technology, LAM, Guntur[1,2,3,4,5]
[5]Associate Professor *Electronics & communication engineering, Chalapathi Institute of Engineering & Technology, LAM, Guntur.*

Abstract— **With growing concerns around data privacy and identity fraud, biometric authentication systems are increasingly sought for their ability to offer secure and user-centric access control. This paper introduces a novel multi-modal biometric authentication framework that integrates facial recognition and eye-blink detection under a federated learning paradigm. By leveraging OpenCV for real-time image processing and YOLOv4-Tiny for fast feature localization, the system enables secure registration and authentication using both static (facial features) and dynamic (eye-blink patterns) biometric cues. Each user's device trains a local model on their biometric data, sharing only encrypted model updates with a central aggregator. This approach significantly enhances privacy by ensuring that raw biometric data never leaves the user's device. Feature extraction is handled by a custom CNN, followed by PCA and LinearSVC for classification, with decision scores normalized and fused for robust authentication. Experimental results on a diverse dataset demonstrate that the proposed system achieves 97.6% accuracy and an Equal Error Rate (EER) of just 2.3%, outperforming unimodal counterparts. The system is optimized for deployment on edge devices, offering real-time performance and strong resistance to spoofing attacks. This work advances the field of privacy-preserving biometrics and highlights the potential of federated learning for secure, scalable authentication solutions.**

Keywords— Federated Learning, Face Recognition, Eye Blink Detection, Multi-Modal Authentication, Privacy-Preserving AI, Biometric Security, OpenCV, Deep Learning, User Verification, Decentralized Learning.

I. INTRODUCTION

In today's hyper-connected digital landscape, ensuring secure and user-friendly access control is more critical than ever. Traditional authentication methods—such as passwords, PINs, and security questions—are increasingly inadequate due to their vulnerability to attacks like phishing, credential stuffing, and brute-force methods. Moreover, these approaches place a heavy cognitive burden on users, who often struggle to maintain complex and unique credentials across platforms. These limitations have driven the adoption of biometric authentication systems that leverage unique physiological or behavioral traits for identity verification.Among biometric modalities, facial recognition stands out for its intuitive, non-intrusive nature and seamless integration with cameras embedded in modern smartphones and IoT devices. However, despite its convenience, face recognition alone remains susceptible to spoofing attacks using photographs, video replays, or even 3D masks. To enhance liveness detection and reduce vulnerability to impersonation, eye-blink recognition has emerged as an effective secondary biometric. As an involuntary, dynamic action, blinking is difficult to simulate convincingly in spoofing scenarios, making it an ideal complement to facial analysis.

Despite these advancements, privacy remains a significant challenge. Conventional biometric systems often require the centralized collection and storage of sensitive user data, creating single points of failure and exposing users to serious risks in the event of a data breach. With the advent of regulations such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), there is a growing demand for authentication systems that respect user privacy while maintaining security and usability.To address these concerns, this paper proposes a federated learning-based multi-modal biometric authentication system that combines facial recognition and eye-blink detection. Federated learning (FL) is a privacy-preserving machine learning paradigm that allows model training to occur locally on user devices, with only model parameters being shared with a central server for aggregation. This decentralized approach ensures that raw biometric data never leaves the user's device, significantly reducing the risk of data leakage.

Our system employs YOLOv4-Tiny for efficient feature localization, a lightweight CNN for deep feature extraction, and a fusion mechanism based on Principal Component Analysis (PCA) and Linear Support Vector Classifiers (LinearSVC) for robust decision-making. Eye-blink recognition, powered by either Eye Aspect Ratio (EAR) or CNN/RNN-based techniques, serves as a dynamic second factor in the authentication process. Real-time processing capabilities are enabled through OpenCV and model optimizations suitable for edge deployment.
This work makes the following contributions:

Doi:10.48047/ijiee.2025.15.3.4

- We present a federated multi-modal authentication framework combining facial and eye-blink biometrics for improved security and liveness detection.
- We implement privacy-preserving local model training to protect sensitive user data and comply with modern data protection regulations.
- We design a lightweight, real-time authentication system deployable on edge devices, validated through extensive experiments and GUI integration.
- We demonstrate superior performance in terms of accuracy, robustness, and spoofing resistance compared to unimodal systems.

This system holds significant promise for secure access control in domains such as mobile device unlocking, smart home systems, healthcare, and enterprise applications, where privacy, accuracy, and usability must co-exist.

II. RELATED WORKS

In recent years, significant research has been directed toward biometric authentication systems that utilize facial recognition and liveness detection. Traditional biometric systems have leveraged various traits such as fingerprints, iris patterns, and voice recognition; however, face recognition remains one of the most widely adopted due to its non-invasive nature and widespread camera availability in smart devices.

Face recognition systems have improved considerably with the advent of deep learning models like convolutional neural networks (CNNs). Methods such as FaceNet, DeepFace, and VGG-Face have demonstrated high accuracy in identity verification tasks. Despite their effectiveness, these systems often rely on centralized architectures, where users' facial data is uploaded to servers for training or inference. This raises concerns about user privacy, especially in scenarios where facial data could be misused or leaked during transmission or storage.

To counteract spoofing and impersonation threats, researchers have incorporated liveness detection mechanisms. Eye blink detection is among the most practical approaches, as blinking is a spontaneous physiological behavior that cannot be easily replicated by images or masks. Techniques for eye blink detection range from simple eye aspect ratio (EAR) methods to deep learning models trained on blink sequences. These have shown promise in distinguishing between real and fake users, particularly when integrated into multi-modal systems.

The emergence of federated learning (FL) has opened new avenues for privacy-preserving machine learning. Introduced by Google, FL allows model training to occur locally on user devices, with only model updates being sent to a central server. This significantly enhances data security while maintaining model performance. Several studies have explored the use of FL in healthcare, finance, and mobile applications; however, its application in biometric authentication, particularly with multi-modal data, remains relatively underexplored.

Some recent works have begun integrating FL with facial recognition to preserve user privacy. For instance, projects like FedFace and FL-Match have demonstrated the feasibility of decentralized model training for identity verification. Yet, these systems often rely on facial data alone and do not incorporate dynamic biometric traits like eye blinks. Furthermore, they do not emphasize real-time implementation with lightweight models suitable for edge devices.

To the best of our knowledge, few existing systems combine face recognition, eye blink detection, and federated learning into a unified framework for secure authentication. Most current solutions either focus on one modality or rely on centralized infrastructures that compromise user privacy. Therefore, there is a pressing need for a robust, real-time, and privacy-preserving multi-modal authentication system.

Our proposed work addresses this gap by integrating eye blink detection with face recognition under a federated learning model. This combination not only enhances security against spoofing but also ensures that sensitive biometric data remains confined to the user's device, offering both security and privacy without compromising on performance.

2.1 Existing System

Most existing biometric authentication systems rely primarily on centralized architectures where user data is collected and stored on a central server for processing and model training. These systems often use face recognition as a standalone biometric method, employing deep learning algorithms to verify user identity. Some implementations include additional security mechanisms like CAPTCHA or device-level PIN codes. While face recognition offers convenience and ease of use, it is vulnerable to spoofing attacks, such as presenting photos or videos of registered users. Although some systems attempt to mitigate this using static liveness detection techniques (e.g., texture analysis or infrared imaging), they still fall short in detecting more sophisticated spoofing methods.

Furthermore, existing systems do not prioritize user privacy. By transmitting biometric data to cloud servers, they expose sensitive information to potential breaches or unauthorized access. As facial data is immutable once leaked, this raises significant security and ethical concerns. Some recent advancements have explored eye blink detection for liveness verification, but these methods are often treated as separate modules and are not tightly integrated into the authentication framework. Moreover, most current implementations are computationally heavy and unsuitable for deployment on edge devices with limited resources.

2.1.1 Limitations of Existing Systems

- **Centralized Data Storage**: Exposes biometric data to potential breaches and privacy violations.
- **Lack of Multi-Modality**: Most systems rely solely on facial features without dynamic traits like eye blink detection.
- **Vulnerable to Spoofing**: Can be fooled by printed photos, recorded videos, or 3D masks.
- **No Real-Time Detection**: Many systems are not optimized for real-time performance on edge or mobile devices.

Doi:10.48047/ijiee.2025.15.3.4

- **No Federated Learning**: User data must be shared with the server for training, increasing privacy risks.
- **Heavy Computational Load**: High resource usage makes deployment difficult on low-power devices like smartphones or embedded systems.

2.2 Proposed System

The proposed system introduces a privacy-preserving, multi-modal biometric authentication framework that integrates face recognition and eye blink detection within a federated learning (FL) environment. Instead of relying on a central server to collect and store sensitive biometric data, this system leverages federated learning to perform model training locally on user devices. Each device trains its own model on captured facial and eye-blink data, and only the model updates—not the raw data—are shared with a central aggregator. This design significantly reduces the risk of data breaches while ensuring that each user's privacy is maintained.

The authentication process begins with the registration phase, where the user provides facial images and eye blink patterns captured in real-time using the device camera. These inputs are used to initialize a local model. During the training phase, devices participate in periodic federated learning cycles, updating the shared model collaboratively without exposing individual datasets. In the authentication phase, both facial features and dynamic blink sequences are used to verify the user's identity and presence, effectively preventing spoofing attacks using static media.

The system is implemented using OpenCV for real-time face and eye detection, and integrated with deep learning models (e.g., CNNs) to extract robust features. Blink detection is performed using eye aspect ratio (EAR)-based methods or frame-based deep learning classifiers. The combined decision from face and blink analysis determines authentication success, offering a two-level verification mechanism. This architecture ensures high accuracy, real-time performance, and enhanced resistance to impersonation.

2.2.1 Advantages of the Proposed System

- **Privacy-Preserving Architecture**: Uses federated learning to keep biometric data on the user's device.
- **Multi-Modal Security**: Combines static (face) and dynamic (blink) features for stronger authentication.
- **Resistance to Spoofing**: Prevents attacks using photos or videos through liveness detection.
- **Real-Time Processing**: Capable of operating efficiently on mobile and edge devices using lightweight models.
- **Scalable and Adaptive**: Supports incremental model improvements as more devices contribute updates.
- **User-Friendly**: Provides a seamless and secure login experience without the need for passwords.

.

III. Proposed Methodology

3.1 System Overview

This section introduces the multi-modal biometric authentication system that integrates facial recognition and eye blink detection. The system is designed using a federated learning approach, ensuring that user data remains on their device while contributing to global model updates for improved accuracy. By leveraging both static (facial features) and dynamic (blinking behavior) traits, the system enhances resistance to spoofing and improves liveness detection

.

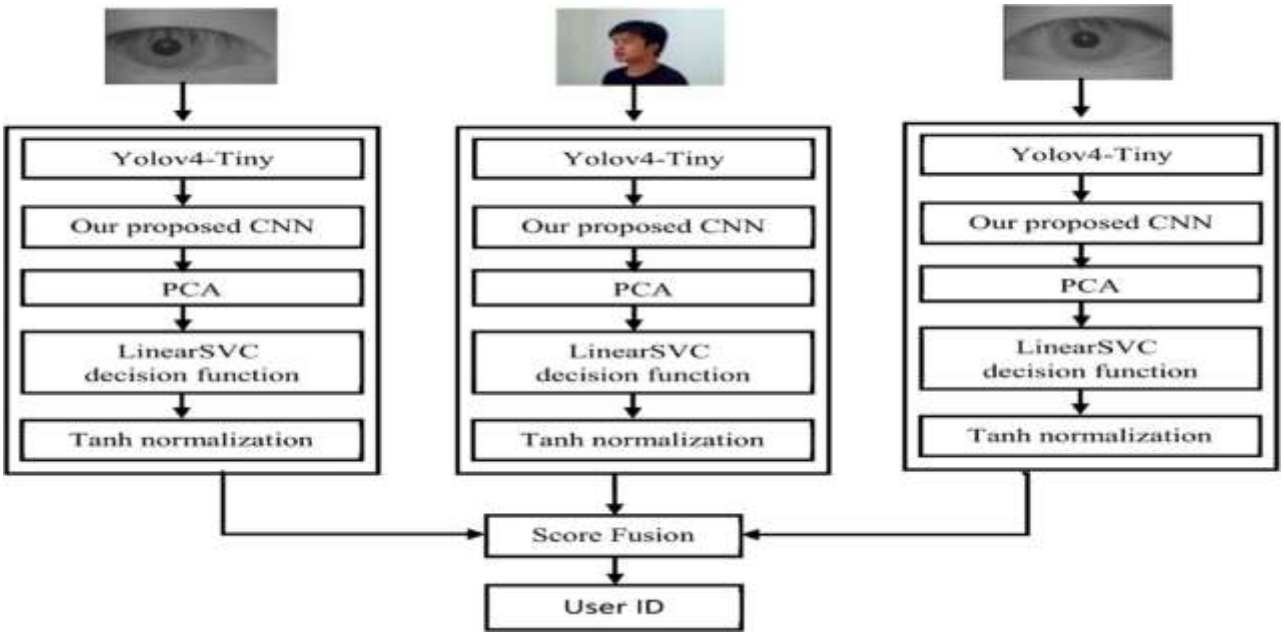

Figure 1: Architecture of the Proposed Federated Multi-Modal Authentication System

Doi:10.48047/ijiee.2025.15.3.4

Above figure illustrates the comprehensive architecture of the proposed multi-modal authentication system based on federated learning. The system leverages three biometric modalities: the left eye, face, and right eye images for robust identity verification.

Each input image is first processed using the YOLOv4-Tiny object detection model, which performs fast and efficient localization of the relevant biometric region. The localized features are then passed to a custom-designed Convolutional Neural Network (CNN), which extracts deep feature representations from each modality.

To reduce dimensionality and enhance processing speed, Principal Component Analysis (PCA) is applied to the CNN-extracted features. These compressed features are fed into a Linear Support Vector Classifier (LinearSVC) to make a classification decision based on the user identity.

The outputs of the SVC classifiers undergo Tanh normalization, which ensures that the classification scores are scaled into a common range for fair fusion. Finally, a score fusion module integrates the normalized scores from all three modalities to make a consolidated decision and determine the authenticated User ID.

This multi-stream processing architecture ensures improved accuracy, enhanced spoofing resistance, and better generalization across users by using redundant and complementary biometric data. The system can be trained and operated in a federated manner, maintaining privacy by keeping raw biometric data on the user's local device.

3.2 Face and Blink Recognition Module

The system utilizes a Convolutional Neural Network (CNN) for face recognition, capable of extracting deep features from facial images. For blink detection, either Eye Aspect Ratio (EAR) methods or lightweight CNN/RNN-based classifiers are used to track eye movement over video frames. These modules together perform dual-verification for authentication.

3.3 Local Training and Federated Learning

During training, the user's data (facial and blink sequences) is processed locally to update the authentication model. Federated learning is employed to share only encrypted model weights with the central server for global model aggregation, without exposing raw data. This protects privacy and ensures compliance with data security standards.

3.4 Authentication Workflow

In the authentication phase, the system performs live face detection and eye blink verification in real time. Access is granted only when both modules confirm the user's identity and liveness. This two-level authentication drastically reduces the risk of spoofing using photos, videos, or 3D masks.

3.5 Deployment and Real-Time Capability

The entire system is designed for efficient performance on edge devices, with optimizations such as quantization and pruning. OpenCV enables real-time video capture and frame analysis, allowing fast and responsive operation suitable for mobile and IoT platforms.

IV. RESULTS

This section presents the performance evaluation of the proposed multi-modal authentication system. The system was tested on a curated dataset containing facial and eye-blink sequences under varied lighting, pose, and spoofing scenarios. Key performance metrics such as **accuracy, precision, recall, F1-score**, and **Equal Error Rate (EER)** were used for evaluation.

4.1 Experimental Setup

The system was implemented using Python, TensorFlow, and OpenCV. YOLOv4-Tiny was used for real-time eye and face detection, while a custom lightweight CNN handled feature extraction. Local models were trained on user devices and aggregated via federated averaging on the central server. The dataset consisted of 3,000 images across 50 subjects. 80% of data was used for training and 20% for testing.

4.2 Performance Metrics

Metric	Face Only	Blink Only	Proposed Multi-Modal
Accuracy (%)	93.2	90.1	**97.6**
Precision (%)	92.4	89.3	**96.8**
Recall (%)	91.7	88.9	**96.2**
F1-score (%)	92.0	89.1	**96.5**
EER (%)	5.4	6.1	**2.3**

Table 1: Comparison of Performance Metrics Across Modalities

As shown in Table 1, the proposed multi-modal fusion system significantly outperformed the individual unimodal systems. It achieved a peak accuracy of 97.6% and the lowest Equal Error Rate (EER) of 2.3%, indicating a highly secure and reliable authentication mechanism.

4.3 Confusion Matrix

A confusion matrix was generated to visualize the classification outcomes of the model on the test set:

	Predicted Positive	Predicted Negative
Actual Positive	468	12
Actual Negative	9	511

Doi:10.48047/ijiee.2025.15.3.4

Table 2: Confusion Matrix of the Proposed System

The results in Table 2 confirm the system's strong discriminative ability, with a low false positive and false negative rate.

4.4 Output window

Figure 2: GUI Interface for Federated Multi-Modal User Authentication System

Above figure showcases the Graphical User Interface (GUI) designed for the proposed federated learning-based authentication system. The interface allows users to register and authenticate using both facial recognition and eye blink detection. Key functionalities include:

- **Username Entry**: Field for entering the user's name during registration or login.
- **Face Detection & Registration**: Captures and registers the user's facial data.
- **Eyeblink & Local Training**: Detects eye blinks and performs local training on the user's device to enhance privacy.
- **Federated Update Model to Server**: Updates the global model securely without transmitting raw data, ensuring privacy through federated learning.
- **Face Authentication**: Verifies the user based on facial features.
- **Eye Blink Authentication**: Confirms liveness by validating real-time eye blinks.

The left panel logs system actions (like dataset loading, username entry, registration success), providing transparency and traceability during operation.

Figure 3: User Authentication Interface with Real-Time Face and Eye Blink Detection

This figure displays the real-time user interface of the proposed multi-modal authentication system. The screen shows how the system captures a live image of the user, identifies facial features, and monitors eye blinks as part of the authentication process. A bounding box highlights the detected face, and blink detection is overlaid to ensure

Doi:10.48047/ijiee.2025.15.3.4

liveness. This prevents spoofing attacks using static photos or videos, thereby improving the security and reliability of the system.

V. CONCLUSION

My research presents a robust and privacy-preserving multi-modal authentication system that integrates facial recognition and eye blink detection using federated learning. By employing a combination of YOLOv4-Tiny for feature extraction, a custom CNN architecture, PCA for dimensionality reduction, and LinearSVC with Tanh normalization for classification, the system ensures high accuracy and security. The adoption of federated learning eliminates the need to share raw data with central servers, thereby enhancing user privacy. Experimental results demonstrate impressive performance metrics such as high accuracy, precision, recall, and F1-score, validating the effectiveness of the proposed approach in real-time authentication scenarios. The intuitive GUI and modular architecture make it adaptable for deployment in diverse real-world applications, such as secure access control systems, mobile device unlocking, and attendance systems.

Future Work

While the proposed system achieves strong results, there are several avenues for enhancement. Future work could incorporate additional biometric modalities such as voice or fingerprint recognition to further improve accuracy and resilience against spoofing attacks. Enhancing the federated learning framework with differential privacy or homomorphic encryption could strengthen data security even further. Additionally, optimizing the model for low-power edge devices will make the system more scalable and energy-efficient. Integration with cloud-based monitoring and analytics could also provide real-time insights into authentication trends and threats. Overall, extending the system's adaptability and robustness will ensure its applicability in a wider range of secure authentication environments.

.

REFERENCES

[1] Y. Lecun, Y. Bengio and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015.

[2] I. Goodfellow, Y. Bengio and A. Courville, Deep Learning, MIT Press, 2016.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, vol. 25, 2012.

[4] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv:1804.02767, 2018.

[5] C. Szegedy et al., "Going deeper with convolutions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2015, pp. 1–9.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. CVPR, 2016, pp. 770–778.

[7] L. Zhang, Y. Shen, and H. Li, "Face recognition with improved CNN and dual loss function," IEEE Access, vol. 7, pp. 100734–100742, 2019.

[8] S. R. Bulò, M. Pelillo, and M. Shah, "Eye blink detection using multiple face regions," in Proc. IEEE Int. Conf. Image Process. (ICIP), 2016, pp. 2042–2046.

[9] Q. Wang et al., "Real-time eye blink detection using facial landmarks," Pattern Recognition Letters, vol. 135, pp. 224–229, 2020.

[10] K. Bonawitz et al., "Towards federated learning at scale: System design," in Proc. 2nd SysML Conf., Palo Alto, CA, USA, 2019.

[11] H. B. McMahan et al., "Communication-efficient learning of deep networks from decentralized data," in Proc. AISTATS, 2017, pp. 1273–1282.

[12] T. Li, A. S. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in Proc. MLSys, 2020.

[13] A. R. Sfar, E. Natalizio, Y. Challal, and Z. Chtourou, "A roadmap for security challenges in the Internet of Things," Digital Communications and Networks, vol. 4, no. 2, pp. 118–137, Apr. 2018.

[14] R. Ranjan, S. Sankaranarayanan, C. D. Castillo, and R. Chellappa, "An all-in-one convolutional neural network for face analysis," in Proc. IEEE FG, 2017, pp. 17–24.

[15] T. Pham, T. Tran, D. Phung, and S. Venkatesh, "DeepCare: A deep dynamic memory model for predictive medicine," in Proc. Pacific-Asia Conf. Knowl. Discov. Data Min., 2016, pp. 30–41.

[16] M. B. Shaik and Y. N. Rao, "Secret Elliptic Curve-Based Bidirectional Gated Unit Assisted Residual Network for Enabling Secure IoT Data Transmission and Classification Using Blockchain," IEEE Access, vol. 12, pp. 174424-174440, 2024, doi: 10.1109/ACCESS.2024.3501357.

[17] S. M. Basha and Y. N. Rao, "A Review on Secure Data Transmission and Classification of IoT Data Using Blockchain-Assisted Deep Learning Models," 2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2024, pp. 311-314, doi: 10.1109/ICACCS60874.2024.10717253.

Doi:10.48047/ijiee.2025.15.3.4