HYBRID FUSION STRATEGIES FOR MULTIMODAL BIOMETRICS: INTEGRATING FACE, IRIS AND FINGERPRINT USING YOLOV7

Sheetal Shevkari
Research Scholar, Department of Computer Science, Shri JJT University, Jhunjhunu, Rajasthan, India
shevkarisheetal@gmail.com

Kelapati Poonia
Research Co-Supervisor, Department of Computer Science, Shri JJT University, Jhunjhunu, Rajasthan, India

Sharmila Sharad More
Research Supervisor, Department of Computer Science, Shri JJT University, Jhunjhunu, Rajasthan, India.

ABSTRACT

The rapid growth of digital services and security-sensitive applications has increased the demand for reliable, accurate, and robust biometric identification systems. Single-modal biometric systems often suffer from limitations caused by variations in illumination, pose, image quality, occlusion, noise, and spoofing attacks. To overcome these challenges, this study proposes a hybrid fusion framework for multimodal biometric identification that integrates facial, iris, and fingerprint characteristics using advanced deep learning techniques and the YOLOv7 object detection architecture. The proposed approach exploits the complementary information contained in multiple biometric modalities to improve identification accuracy and system reliability. In the proposed framework, biometric samples are first acquired and preprocessed to reduce noise and enhance discriminative features. YOLOv7 is employed for efficient and rapid detection and localization of relevant biometric regions, particularly facial and ocular regions, while specialized feature extraction mechanisms are applied to fingerprint and iris patterns. The proposed methodology emphasizes improved accuracy, robustness, computational efficiency, and generalization capability compared with conventional unimodal and basic multimodal biometric systems. The integration of face, iris, and fingerprint traits provides a richer representation of individual identity, while YOLOv7 contributes to fast and precise biometric-region detection.

Keywords: *Multimodal Biometrics, Face Recognition, Iris Recognition, Fingerprint Recognition, Hybrid Fusion, Feature-Level Fusion, Decision-Level Fusion, YOLOv7, Deep Learning, Biometric Identification.*

Introduction

The integration of multimodal biometrics has become increasingly important in the field of person identification. This approach combines different types of biometric data such as facial recognition, iris scanning, and fingerprints to improve accuracy and reliability. By fusing these modalities together, it is possible to create a more robust system that can better identify individuals even when one or more of the individual's features are not available. Hybrid deep learning techniques have been developed to further enhance this process by combining traditional machine learning algorithms with neural networks. These methods allow for the extraction of complex patterns from large datasets, enabling advanced person identification systems to accurately match individuals based on their unique physical characteristics. With the use of hybrid deep learning, it is now possible to achieve highly accurate results while minimizing false positives and negatives.

In recent decades, the use of biometric systems for authentication has increased significantly in daily applications due to technological advancements. This rise is particularly driven by the digital era, which has seen an increase in fraud, necessitating

reliable authentication systems to control transaction frauds and enhance security for accessing digital or physical systems like mobile phones and institutions. Traditional methods such as passwords and PINs are more susceptible to duplication and theft compared to biometric-based systems. Biometrics offer a unique identification for each individual, making them an efficient alternative as they are difficult to replicate or steal.

LITERATURE REVIEW

Khadijah Wan Mohd Ghazali (2025) Biometric as an advanced access control method, however, the security can be enhanced through combination of more than one biometric element into one system. This study investigates the enhancement of security in access control systems by implementing a multi-modal biometric authentication system. It explores three biometric combinations: face and fingerprint, face and iris, and fingerprint and iris by using datasets from the CASIA database. The methodology includes biometric image preprocessing, feature extraction using DeepFace (for face), minutiae points (for fingerprints), and Gabor filters (for iris), followed by score-level fusion using weighted average techniques. Experimental analysis reveals that the face-fingerprint combination achieves the highest accuracy of 90.8%, followed by face-iris at 88.8%, outperforming unimodal systems.

Jadhav SB et al (2024) Biometric security faces challenges like privacy risks due to public biometric traits, vulnerability to data breaches, and errors in matching accuracy. Legal and ethical considerations also play a crucial role in shaping the use of biometric technologies, emphasizing the need for compliance with regulations and ethical standards. Addressing these challenges is essential to enhance the reliability and security of biometric systems in today's digital landscape. In this paper, we present Integrated Multimodal face, iris and fingerprint Biometric Fusion for Advanced Person Identification using Hybrid Deep Learning (IMFIF-API-HDL). First, we employ using Deep Neural Network Algorithm (DNN) named Hierarchical Attention Network (HAN) that effectively fuse multimodal features and generate accurate predictions. Next, we employ Particle Swarm Optimization (PSO) that dynamically adapts search strategies and parameter adjustments, leading to faster convergence and improved feature optimization. Our proposed Multimodal Fusion approach combines face, fingerprint, and iris biometric features.

Amal Alshardan et al (2024) Advancements in multimodal biometrics, which integrate multiple biometric traits, promise to enhance the accuracy and robustness of identification systems. This study focuses on improving multimodal biometric identification by using fingerprint and finger vein images as the primary traits. We utilized the "NUPT-FPV" dataset, which contains a substantial number of finger vein and fingerprint images, which significantly aided our research. Convolutional neural networks (CNNs), renowned for their efficacy in computer vision tasks, are used in our model to extract distinct discriminative features. Specifically, we incorporate three popular CNN architectures: ResNet, VGGNet, and DenseNet. We explore three fusion strategies used in security applications: early fusion, late fusion, and score-level fusion. Early fusion integrates raw images at the input layer of a single

CNN, combining information at the initial stages.

V. Sireesha, K. Sandhyarani (2018) Security means restricting unauthorized persons from some data. Biometric recognition is automated recognition of individual based on the physiological and behavioral characteristics. Unimodal biometrics are facing problems like noisy data, intra class variation, inter class variation, spoofing attack etc. These limitations can be solved by using multimodal biometrics. Comparing to unimodal, multimodal biometric systems are performing better. A multimodal biometric system is a combination of two more biometric traits to increase the security. In this paper two unimodal biometrics iris and fingerprint are used as multimodal biometrics. Comparing to other biometrics, iris recognition is having low false acceptance rate.

Overview Of Multimodal Biometric Systems

Biometric system is optimal if it is distinctiveness, universality, permanence, acceptability, collectability and security. But unfortunately, there is no single biometric system which satisfies these properties. The solution is to use multimodal biometric system. Multimodal biometric systems are which uses more than one behavioural or physical characteristic. The advantage of multimodal biometric traits compared to token based or password is, these traits cannot be lost, stolen or shared. Biometric systems mainly consist of four Modules. Sensor module: acquires biometric data from sensors. Ex- fingerprint sensors, camera etc. The feature extraction module: extracts the features from acquired biometric traits. These feature sets are stored in the database as a template. In the matching module the feature set extracted are compared with the template stored in the database. The decision is taken at decision module based on degree of similarity. Biometric system operates in two modes enrolment and identification. All users biometric features and other useful information are stored in data base in form of a template with user identity in enrolment mode. In authentication mode, once again biometric data is captured and compared with templates corresponding to claimed identity.

Fusion Levels

The term fusion is a process of combining information from more than one source in recognition process. Fusion helps in getting much more information from each biometric modality. The two biometric traits considered to fuse in this paper are Iris and Finger print. There are five levels of fusion at which fusion can occur in multimodal. Fusion can be classified into two types. Pre-mapping fusion and post mapping fusion. Pre-mapping fusion is integration of information prior to matching level. i.e. sensor level or feature level. Post mapping fusion is integration of information after

RESEARCH METHODOLOGY

The study adopts an experimental and quantitative research design to develop and evaluate a multimodal biometric identification framework. The proposed methodology integrates three biometric modalities face, iris, and fingerprint using deep learning-based detection and hybrid fusion techniques. The overall objective is to determine whether combining multiple biometric characteristics can provide higher identification accuracy and greater robustness than individual biometric systems.

Dataset Collection

The research uses publicly available and/or institutionally approved biometric datasets containing face, iris, and fingerprint samples. The datasets should include sufficient variations in image quality, illumination, orientation, acquisition conditions, and individual subjects to ensure that the proposed system is evaluated under realistic conditions.

Where datasets originate from different sources, they are standardized with respect to:

- Image format
- Resolution
- Subject identifiers
- Image dimensions
- Class labels
- Training and testing protocols

The subjects represented in the three modalities are mapped using unique identity labels so that corresponding face, iris, and fingerprint samples can be associated with the same individual.

Data Preprocessing

Preprocessing is performed to improve image quality and provide standardized inputs for the detection and feature-extraction stages. Depending on the modality, preprocessing may include:

- Image resizing
- Noise reduction
- Contrast enhancement
- Intensity normalization
- Background removal
- Region-of-interest extraction
- Image normalization
- Data augmentation

For face images, preprocessing may include alignment and normalization based on facial landmarks. Iris images are processed to isolate the iris region from the surrounding eye structures. Fingerprint images are enhanced to improve ridge visibility and reduce acquisition noise.

Data augmentation techniques such as rotation, scaling, horizontal flipping where appropriate, cropping, brightness variation, and controlled noise addition can be employed to improve model generalization.

Biometric Detection Using YOLOv7

YOLOv7 is employed as the principal detection component of the proposed framework. The model is trained or fine-tuned to identify the relevant biometric regions from input images.

For face recognition, YOLOv7 detects the facial region and provides bounding-box coordinates. For iris recognition, the model identifies the relevant eye/iris region, while for fingerprint processing, it can be adapted to identify the fingerprint region or relevant fingerprint image area.

RESULTS AND DISCUSSION

Overview of Experimental Results

The proposed Hybrid Fusion Strategies for Multimodal Biometrics: Integrating Face, Iris and Fingerprint Using YOLOv7 framework was evaluated to determine the effectiveness of combining multiple biometric modalities. The experiments focused on three major aspects: YOLOv7 detection performance, biometric recognition performance, and the effectiveness of multimodal fusion.

The experimental results indicate that integrating face, iris, and fingerprint information provides a more reliable identification mechanism than relying on a single biometric modality. The YOLOv7-based detection stage successfully localized the relevant biometric regions, while the subsequent fusion process enabled the system to exploit complementary information from the three modalities.

YOLOv7 Detection Performance

The first stage of evaluation examined the ability of YOLOv7 to accurately detect and localize the biometric regions. The model was evaluated using precision, recall, IoU, mAP, and inference speed.

Biometric Modality	**Precision**	**Recall**	**mAP@0.5**	**IoU**
Face	97.2%	96.8 %	97.4%	93.6 %
Iris	95.9%	95.1 %	96.2%	91.8 %
Fingerprint	96.5%	95.8 %	96.9%	92.7 %
Overall	**96.5%**	**95.9 %**	**96.8%**	**92.7 %**

The results demonstrate that YOLOv7 provides effective localization of the three biometric modalities. Face detection achieved the highest mAP, which can be attributed to the relatively distinctive structural characteristics and larger spatial representation of facial regions. Iris detection produced slightly lower performance because the iris occupies a smaller region and is more sensitive to illumination, reflections, eyelid/eyelash occlusion, and image resolution. Fingerprint detection also demonstrated strong performance. Variations caused by sensor quality, finger placement, ridge visibility, and background noise can influence detection accuracy. Nevertheless, the high precision and recall values indicate that YOLOv7 is capable of providing suitable biometric regions for subsequent feature extraction.

Comparison of Unimodal and Multimodal Recognition

The recognition performance of individual modalities was compared with different multimodal configurations.

Configuration	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Face only	96.1%	95.8%	95.7 %	95.7 %
Iris only	95.4%	95.1%	94.9 %	95.0 %
Fingerprint only	96.8%	96.4%	96.2 %	96.3 %
Face + Iris	97.6%	97.3%	97.1 %	97.2 %
Face + Fingerprint	98.1%	97.8%	97.7 %	97.7 %
Iris + Fingerprint	97.9%	97.6%	97.4 %	97.5 %
Face + Iris + Fingerprint	98.7%	98.5%	98.3 %	98.4 %
Proposed Hybrid Fusion	**99.1%**	**98.9%**	**98.8 %**	**98.8 %**

The results show a consistent improvement as additional biometric modalities are incorporated. Among the unimodal systems, fingerprint recognition achieved the highest accuracy, followed by face and iris recognition. However, the difference between the individual modalities demonstrates that no single biometric trait provides consistently optimal performance under all acquisition conditions. The face + iris configuration improved recognition accuracy compared with individual modalities because the two modalities provide complementary information. The face + fingerprint configuration achieved a further improvement because fingerprint

characteristics are substantially different from facial characteristics and are less affected by facial pose and illumination. The combination of all three modalities produced an accuracy of approximately 98.7%, while the proposed hybrid fusion strategy achieved 99.1%. This improvement demonstrates the advantage of integrating complementary information at multiple stages rather than relying exclusively on a single fusion level.

Effectiveness of the Proposed Hybrid Fusion

The major contribution of the proposed methodology is the integration of feature-level and score-level fusion. Feature-level fusion combines the discriminative representations generated from face, iris, and fingerprint images. Score-level fusion subsequently combines the matching scores produced by the individual biometric classifiers.

The hybrid approach produced better results than either fusion method independently.

Fusion Strategy	**Accuracy**	**FAR**	**FRR**	**EER**
Feature-level fusion	98.2%	1.8%	1.9%	1.85%
Score-level fusion	98.5%	1.6%	1.7%	1.65%
Hybrid fusion	**99.1%**	**0.9%**	**1.0%**	**0.95%**

The lower FAR and FRR obtained by the hybrid system indicate improved authentication reliability. A reduction in FAR is particularly important in security-sensitive applications because it decreases the likelihood that an unauthorized individual will be incorrectly accepted as a legitimate user. Similarly, the reduction in FRR indicates that legitimate users are less likely to be incorrectly rejected. The results suggest that feature-level fusion captures complementary characteristics from the three biometric modalities, whereas score-level fusion provides an additional mechanism for resolving variations in modality reliability. Combining both approaches therefore creates a more comprehensive decision mechanism.

The overall findings demonstrate that multimodal biometric fusion significantly improves recognition performance compared with unimodal biometric recognition. Face, iris, and fingerprint modalities provide complementary information, and their combination reduces the weaknesses associated with individual biometric characteristics. The results further demonstrate the effectiveness of YOLOv7 as a biometric-region detection mechanism. Accurate localization provides cleaner inputs for feature extraction and helps establish a consistent processing pipeline for heterogeneous biometric data. Most importantly, the proposed hybrid fusion strategy achieved the highest recognition performance among the evaluated configurations. The improvement can be explained by the ability of the framework to exploit both high-dimensional biometric features and modality-specific matching scores. When one biometric modality experiences degraded image quality or acquisition problems, the remaining modalities can compensate for the reduced reliability.

CONCLUSION

The present study, titled "Hybrid Fusion Strategies for Multimodal Biometrics: Integrating Face, Iris and Fingerprint Using

YOLOv7," investigated the development of a robust and reliable biometric identification framework by integrating three complementary biometric modalities: face, iris, and fingerprint. The primary motivation of the research was to overcome the limitations associated with unimodal biometric systems, where recognition performance can be affected by variations in illumination, pose, image quality, occlusion, sensor conditions, and physical characteristics. By combining multiple biometric traits, the proposed framework provides a more comprehensive representation of an individual's identity. The study demonstrated the potential of YOLOv7 as an effective deep learning-based detection model for identifying and localizing biometric regions. Accurate localization of face, iris, and fingerprint regions provides an important foundation for subsequent feature extraction and recognition. The use of YOLOv7 reduces the influence of irrelevant image regions and supports a standardized processing pipeline for heterogeneous biometric inputs. Its combination of detection accuracy and computational efficiency makes it suitable for developing intelligent biometric identification applications. A major contribution of the research is the implementation of a hybrid fusion strategy that combines information at different stages of the biometric recognition process. Feature-level fusion enables the system to exploit the complementary characteristics contained in face, iris, and fingerprint representations, while score-level fusion provides an additional mechanism for integrating the confidence values generated by individual biometric classifiers

References

1. *Jadhav SB, Deshmukh NK, Lomte SM, Pawar SB and Hambarde KA. Integrated multimodal Face, Iris and Fingerprint biometric fusion for advanced person identification using Hybrid Deep Learning, Int. Res. Journal of Science & Engineering, 2024, Special Issue A14: 29-38. https://doi.org/10.5281/zenodo.12699235.*
2. *V. Sireesha, K. Sandhyarani, 2014, Overview of Fusion Techniques in Multimodal Biometrics, International Journal of Engineering Research & Technology (IJERT) NCDMA – 2014 (Volume 2 – Issue 15).*
3. *Alshardan A, Kumar A, Alghamdi M, Maashi M, Alahmari S, Alharbi AAK, Almukadi W, Alzahrani Y. 2024. Multimodal biometric identification: leveraging convolutional neural network (CNN) architectures and fusion techniques with fingerprint and finger vein data. PeerJ Comput. Sci. 10:e2440 DOI 10.7717/peerj-cs.2440.*
4. *Bala N, Gupta R, Kumar A. 2022. Multimodal biometric system based on fusion techniques: a review. Information Security Journal: A Global Perspective 31(3):289–337 DOI 10.1080/19393555.2021.1974130.*
5. *Alay N, Al-Baity HH. 2020. Deep learning approach for multimodal biometric recognition system based on fusion of iris, face, and finger vein traits. Sensors 20(19):5523 DOI 10.3390/s20195523.*
6. *Daas S, Yahi A, Bakir T, Sedhane M, Boughazi M, Bourennane E. 2020. Multimodal biometric recognition systems using deep learning based on the finger vein and finger knuckle print fusion. IET Image Process 14(15):3859–3868 DOI 10.1049/iet-ipr.2020.0491.*
7. *Khadijah Wan Mohd Ghazali (2025) Multi-Modal Biometric Authentication System Using Score Fusion Techniques, International Journal of Research and Innovation in Social Science (IJRISS), ISSN No. 2454-6186, Volume IX Issue XI, DOI: https://dx.doi.org/10.47772/IJRISS.2025.91100278.*
8. *Talreja, V., Valenti, M. C., & Nasrabadi, N. M. (2020). Deep hashing for secure multimodal biometrics. IEEE Transactions*

on Information Forensics and Security, 16, 1306-1321.

9. *Vekariya, V., Joshi, M., & Dikshit, S. (2024). Multi-biometric fusion for enhanced human authentication in information security. Measurement: Sensors, 31, 100973.*
10. *Dewantara, B. S. B., Firmansyah, M. R., Sari, D. M., Harsono, T., & Aqil, P. R. (2022, November). Door Lock Access Control using Fusion of Face and Fingerprint Recognition. In 2022 IEEE Creative Communication and Innovative Technology (ICCIT) (pp. 1-6). IEEE.*