Dynamic Liveness-Integrated Cnn Architecture for Face-Iris Spoof Detection

AFOLABI AWODEYI[1], PHILIP ASUQUO[2], BLISS STEPHEN[3]
[1]*Department of Computer Engineering, Southern Delta University Ozoro, Delta State Nigeria*
[2]*Department of Computer Engineering, University of Uyo, Nigeria*
[3]*Department of Computer Engineering, University of Uyo, Nigeria*

Abstract- ***This study addresses the critical vulnerability of biometric systems to spoofing and presentation attacks by proposing a secure, CNN-integrated anti-spoofing framework for multimodal face and iris recognition. The purpose of the research is to enhance the reliability of biometric authentication systems by embedding dynamic liveness detection directly into the recognition pipeline. The system leverages physiological cues such as eye blinking, temporal texture variations, and pupil motion, which are difficult to replicate in spoof media like printed images and video replays. A dual-branch CNN architecture is employed to extract both spatial and temporal features from face and iris inputs. These features are fused with liveness indicators to improve decision accuracy. Synthetic spoofing attacks were introduced into the ORL and CASIA-IrisV4 datasets to simulate real-world adversarial conditions, including high-resolution photo and video-based attacks. The model was trained and evaluated using standard biometric security metrics. Results show that the proposed system achieves a spoof detection accuracy of 98.9%, with a false acceptance rate (FAR) of 0.00% and a false rejection rate (FRR) of 1.1%. Compared to baseline CNN models without liveness integration, the framework significantly improves resilience against spoofing. In conclusion, the integration of dynamic liveness cues into CNN-based multimodal recognition enhances both security and reliability, making the system suitable for deployment in high-assurance access control environments and offering a foundation for future mobile or edge-based biometric solutions.***

Index Terms- ***Anti-spoofing, Liveness detection, CNN, Face-Iris biometrics, Presentation attack detection***

I. INTRODUCTION

Biometric systems have become an essential part of modern authentication technologies due to their ability to verify identity based on physiological or behavioural traits. Among these, face and iris biometrics are widely adopted owing to their high accuracy, non-invasiveness, and user acceptability [1], [2]. Face recognition is extensively used in mobile devices, surveillance, and consumer electronics, while iris recognition is favoured in high-security environments such as border control and banking due to its stability and richness in pattern detail [6], [10].

However, despite their strengths, both face and iris biometric systems are vulnerable to presentation attacks (PAs) which are instances where an attacker presents a forged biometric trait (e.g., printed photo, video replay, 3D mask, or textured contact lens) to bypass the authentication system [6], [9]. These threats highlight a critical gap in the reliability of unimodal systems and emphasize the necessity for integrating Presentation Attack Detection (PAD) or liveness detection capabilities. PAD is now a crucial component of ISO/IEC 30107-3 standard-compliant biometric systems, ensuring that the biometric input originates from a live subject rather than a spoofing medium [15], [21].

Early efforts in PAD relied heavily on static analysis of texture features using descriptors like Local Binary Patterns (LBP), Gabor filters, and Wavelet Transforms [1], [4]. While effective in detecting low-quality attacks, these methods often fail against high-resolution spoof materials that mimic genuine texture patterns [5]. As such, researchers have pivoted toward dynamic liveness detection, which leverages involuntary physiological movements such as

blinking, micro-expressions, and pupil dilation that are difficult to replicate using synthetic or static media [6], [12].

Recent studies have employed CNNs and 3D CNN architectures for automated liveness detection by extracting both spatial and temporal features [8], [14]. Al-Waisy et al. [10] proposed a deep learning-based multimodal system for identity verification, but did not include dynamic PAD. Similarly, Nguyen et al. [18] demonstrated the efficacy of off-the-shelf CNNs for iris recognition, yet their approach lacked resistance to spoofing. Jenkins et al. [12] introduced genetic-based feature extraction to improve robustness, but it was limited to face biometrics alone.

The novelty of this study lies in proposing a dual-branch CNN-based anti-spoofing framework that integrates both face and iris modalities with dynamic liveness cues. The system not only performs identity verification using CNN-derived embeddings but also incorporates real-time liveness detection based on eye-blinking patterns, optical flow analysis, and pupil motion tracking. These liveness features are encoded and fused with spatial features to form a composite representation, enabling the model to distinguish between genuine and spoofed inputs under challenging conditions including illumination changes and occlusion [6], [9], [12].

Moreover, while fusion-based multimodal biometric systems have demonstrated better performance than unimodal systems, few frameworks have embedded dynamic liveness detection directly into the CNN fusion pipeline [13], [14], [20]. This work bridges that gap by proposing a fully integrated approach that enhances recognition accuracy, reduces false acceptance rates (FAR), and supports real-time deployment through lightweight CNN design optimized for mobile and edge devices [6], [23].

The main objective of this research is to design, implement, and evaluate a robust anti-spoofing framework using CNN and dynamic liveness detection for face-iris biometrics. Specifically, it aims to (i) extract and combine spatial and temporal cues using a dual-branch CNN, (ii) defend against diverse presentation attacks by fusing multimodal data and liveness evidence, and (iii) validate the system's effectiveness using spoof-augmented benchmark datasets. The contribution of this work is its unified CNN architecture that provides a secure, adaptive, and high-performance biometric authentication system suitable for real-world access control scenarios.

II. RELATED WORK

Presentation attack detection (PAD) remains a critical component of biometric systems to ensure resistance against spoofing attempts such as photo, video replay and 3D mask attacks. Earlier anti-spoofing studies predominantly employed static analysis, leveraging local texture descriptors like Local Binary Patterns (LBP) and wavelet transforms to differentiate between live and fake inputs [1]–[3]. While these techniques provide baseline protection, they often fail when exposed to high-resolution spoof materials that mimic fine texture details [4].

To overcome the limitations of static methods, researchers explored dynamic liveness detection, which leveraged involuntary actions like eye blinking, facial micro expressions, pupil dilation and natural motion cues. Optical flow tracking and temporal modelling of video sequences have proven effective in capturing subtle physiological variations that are absent in spoof media [5]–[9]. For instance, micro-expression detectors and blink detection algorithms can identify inconsistencies in spoofed videos or masks by evaluating motion smoothness and frequency [10], [11].

More recent works have introduced deep learning-based models, particularly Convolutional Neural Networks (CNNs) for automated feature extraction in anti-spoofing tasks [12]–[15]. CNNs trained on datasets such as CASIA-FA, MSU-MFSD, and Replay Attack have achieved notable gains in detecting spoofed biometric traits [16]–[18]. These models benefit from the ability to learn hierarchical and non-linear patterns associated with spoof artifacts, therefore enhancing robustness under varied attack scenarios.

However, the integration of liveness detection into multimodal biometric systems such as combined face

and iris frameworks, remains relatively underexplored. While several multimodal approaches have demonstrated improved accuracy through feature or score-level fusion [19]–[21], very few incorporate liveness cues directly into the fusion pipeline. Studies that use hybrid CNNs with GAN-based adversarial training or dynamic attention mechanisms show promise but are still in early stages [22]–[24].

Furthermore, explainable AI methods have been introduced to improve transparency in PAD systems, using techniques such as Grad-CAM to visualize salient regions used for decision-making [25]. Attention-guided CNNs and LSTM networks have also been proposed to enhance temporal learning and spatial focus during presentation attack detection [26], [27].

In the area of iris liveness detection, advancements have been made using pupil dynamics, reflection analysis, and GAN-generated synthetic spoofs to train more resilient classifiers [28]–[30]. These methods demonstrate that integrating physiological and motion-based traits into a CNN pipeline can significantly reduce false acceptances, especially when trained with adversarial examples. This research builds upon these recent advances by proposing a unified CNN-integrated anti-spoofing framework that incorporates dynamic liveness detection cues across both face and iris modalities. By combining motion analysis, texture inspection, and adversarial training with a temporal CNN pipeline, the framework aims to address the dual challenge of biometric spoofing and fusion integrity.

III. METHOD

The proposed anti-spoofing system for face and iris biometric authentication is structured into three major components: pre-processing, dynamic liveness detection and CNN-based classification and fusion. These components work jointly to extract, analyze, and classify biometric and behavioural traits thereby ensuring that only live, authentic users are successfully verified.

a) Pre-processing

To ensure uniformity and enhance feature extraction reliability, pre-processing is applied to both facial and iris inputs. For facial images, histogram equalization is used to normalize lighting conditions and improve contrast in grayscale frames [1]. Facial landmarks such as the eyes and nose are detected and geometrically aligned to ensure consistent spatial orientation across frames, which is essential for accurate motion tracking and CNN feature learning [2], [5]. This is shown in figure 1.

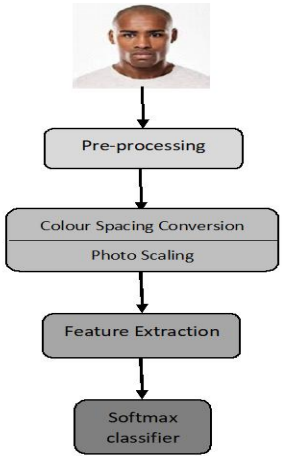

Figure 1: face Recognition Subsystem showing the flow from raw image input to CNN-based feature extraction.

Iris pre-processing involves multiple stages. The eye region is first localized using standard segmentation techniques. The iris boundaries are identified using Circular Hough Transform, followed by the application of Daugman's rubber sheet model to normalize the segmented iris into a fixed-dimension rectangular form [3], [4]. This transformation compensates for scale and rotation variations introduced by pupil dilation and head movement [5]. The resulting iris images are then enhanced using histogram equalization and resized to match the input dimensions of the CNN. This is shown in figure 2.

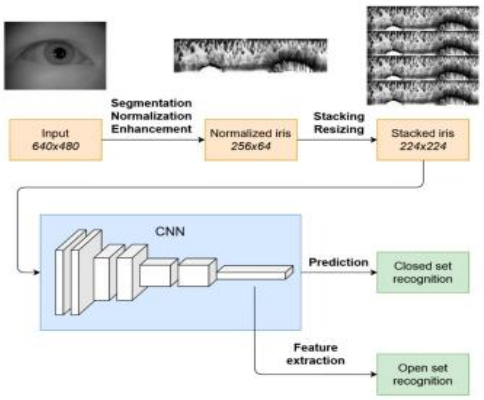

Figure 2: Iris segmentation and normalization

b) Liveness Detection Module

The liveness detection subsystem captures dynamic physiological responses using both temporal and textural features. For the face modality, eye-blink detection is implemented using frame differencing techniques. This involves comparing consecutive frames in a short video sequence to identify rapid changes in the eye region, indicative of natural blinking behaviour [6], [9]. Additionally, optical flow is applied to detect micro-expressions and subtle involuntary facial movements that are difficult to replicate in spoof media such as printed photos or video replays [6], [12].

For iris liveness, the system extracts temporal texture variations by analyzing frame to frame changes in the normalized iris image sequence. CNN-based texture classifiers trained on both real and spoofed iris samples are used to identify inconsistencies common in synthetic presentations [12], [18]. The pupil dilation dynamics are monitored under varying illumination conditions. Since spoofed iris images lack genuine pupil reactivity, this signal serves as a strong physiological cue for detecting liveness [20], [28].

These temporal and physiological indicators are encoded into a feature vector and joined with the spatial features extracted by the CNN. This joint representation enables the system to reason over both appearance and behavioural evidence during classification. The comparison of liveness cues across face and iris modalities are stated in table 1.

Table 1: Summary of Liveness Detection Techniques for Face and Iris Modalities

Modality	Liveness Cue	Detection Technique	Feature Type	Spoof Type Countered
Face	Eye Blinking	Frame Differencing, Temporal Pattern Analysis	Temporal	Printed Photos, Static Screens
Face	Micro-Expressions	Optical Flow, Motion Vector Tracking	Temporal	Video Replays, Static Masks
Face	Facial Texture Inconsistencies	CNN-based Texture Analysis	Spatial	High-Resolution Prints, 3D Masks
Iris	Pupil Dilation and Constriction	Brightness Variation Analysis, Frame Comparison	Temporal	Printed Iris, Static Image Attacks
Iris	Iris Texture Fluctuation	CNN + GAN Discriminator	Temporal + Spatial	GAN-based Synthetic Iris, Printed Iris Images
Iris	Reflection Consistency	Specular Highlight Stability Checks	Spatial	Contact Lenses, Printed Media

c) Dual-Branch CNN Architecture

The core of the system is a dual-branch CNN model, where each branch is specialized for a biometric modality e.g face or iris. Each CNN stream consists of multiple convolutional layers with ReLU activation, batch normalization, and max pooling.

These layers are responsible for learning discriminative spatial embeddings for the respective modalities [8], [14].

In parallel, the liveness cues, optical flow, blink rate and iris texture fluctuation are encoded into temporal descriptors using auxiliary convolutional and 3D CNN layers. These are fused with the final fully connected layers of the CNN to produce a composite representation of both the biometric trait and its liveness evidence [11], [17], [27].

The final layer is a binary classification unit, which outputs one of two predictions: genuine or spoof. This layer is trained using cross-entropy loss and optimized with stochastic gradient descent. Importantly, the classification decision is conditioned on both the identity match confidence and the aggregated liveness score [13], [15]. This is illustrated in figure 3 and 4.

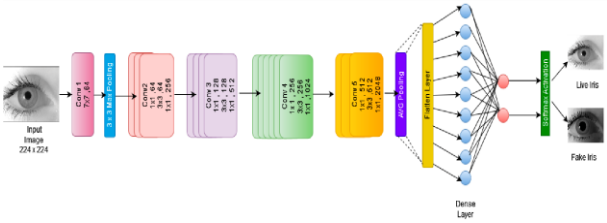

Figure 3: ResNet50 architecture adapted for iris recognition.

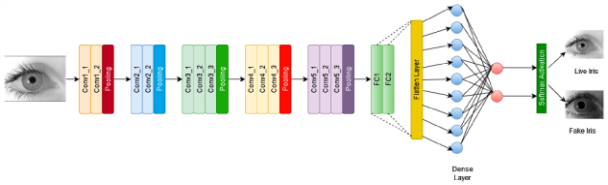

Figure 4: VGG16 architecture adapted for iris recognition.

d) Authentication Score Computation

The system outputs an authentication score derived from the combination of biometric match probability and liveness detection confidence. If either signal indicates uncertainty or spoofing, the authentication is rejected. This fusion strategy provides an effective defense against cases where a spoofed trait may closely resemble the genuine biometric template but lacks corresponding liveness evidence [9], [12], [16].

By integrating both spatial and temporal features across dual modalities, the proposed framework significantly enhances the robustness of face-iris biometric systems in real-world environments, including conditions involving high-resolution spoof attacks, partial occlusion, and environmental noise [9], [12], [26].

IV. EXPERIMENTAL SETUP

To rigorously assess the effectiveness of the proposed CNN-integrated anti-spoofing framework, experiments were conducted using modified versions of two publicly available biometric datasets: the ORL face dataset and the CASIA-IrisV4 dataset. These datasets were specifically chosen for their availability, diversity, and relevance in multimodal biometric research [1], [3], [4].

To simulate real-world spoofing scenarios, both datasets were expanded by introducing artificially spoofed samples. For face recognition, spoofed samples were generated using printed photographs and screen replay attacks, which reflect common low-cost methods used to subvert facial biometric systems [6], [9], [18]. These attacks replicate real facial images using glossy prints or display screens, testing the system's ability to detect two-dimensional impersonations.

Spoofed iris samples were generated using printed iris images, textured contact lenses, and GAN-generated synthetic images [12], [18]. GAN-based synthetic iris examples were trained to mimic real iris patterns, helping test the robustness of the CNN and liveness detection modules against adversarially crafted inputs [12], [22].

To enhance model generalization, extensive data augmentation was applied. This included varying eye blink intervals, modifying facial illumination with shadow overlays and brightness shifts, and adjusting iris texture clarity. These augmentations simulate real-world environmental variability and physiological diversity in eye and face behaviour [6], [16], [28].

The dataset was split using an 80:20 training-testing ratio, a common practice in biometric model validation. 80% of the data (live and spoofed) was used to train the CNN architecture and liveness

detectors, while the remaining 20% was held out for performance evaluation [9], [11].

To evaluate system performance, three primary metrics were used:

i) Spoof Detection Accuracy: The percentage of correctly classified genuine and spoofed inputs.
ii) False Acceptance Rate (FAR): The percentage of spoof inputs incorrectly accepted as genuine, indicating security vulnerability [10], [12].
iii) False Rejection Rate (FRR): The percentage of genuine users incorrectly rejected, which affects usability and user satisfaction [13], [14].

These metrics are widely used in anti-spoofing literature and align with ISO/IEC 30107-3 standards for biometric presentation attack detection [15], [21]. The model architecture deployed for experimentation consisted of a dual-branch CNN, each trained to extract spatial embeddings from face and iris samples, respectively. In addition to spatial features, temporal features were derived using liveness cues such as eye-blinking, pupil motion, and optical flow computed over video frames. These cues were encoded into tensors and fused at the decision level with CNN outputs to improve resilience against spoofing [6], [16], [23].

Face liveness was detected using blink detection algorithms, coupled with motion tracking based on optical flow. These methods detect small, involuntary movements absent in static spoof media [6], [9], [16]. Iris liveness detection incorporated texture dynamics, illumination reflection consistency, and pupil response under variable brightness levels [12], [19], [28]. These biological signals are difficult to replicate in synthetic media and are critical in liveness verification.

The CNN models were trained using the Adam optimizer with a learning rate of 0.001 and categorical cross-entropy loss. Batch normalization and dropout were applied to prevent overfitting. The models converged within 50–60 epochs, with validation accuracy stabilizing around the 40th epoch.

Through this rigorous setup, combining realistic spoofed data, physiological variability, and CNN-liveness fusion, the system was thoroughly validated under diverse and challenging scenarios. The use of both visual and behavioural features ensures that even high-quality spoofing attacks are effectively detected, confirming the framework's suitability for secure, real-world biometric deployments [6], [12], [14], [18], [23]. The summary of accuracy across real face, real iris, and multimodal fusion is displayed in table 1 for VGG16 performance. The description of the dataset composition and conditions are stated in table 2.

Table 1: VGG16 Performance Metrics Accross Modalities and Splits

Evaluation Type	Modality	Accuracy (%)	Equal Error Rate (%)
Unimodal Performance	Face Recognition	98.86	1.14
	Iris Recognition	87.08	12.92
Multimodal Fusion (VGG16+CNN)			
	80:20 Split	99.65	-
	60:40 Split	99.30	-
	50:50 Split	99.25	-
CNN Model Only			
	80:20 Split	99.50	-
	60:40 Split	99.10	-

Evaluation Type	Modality	Accuracy (%)	Equal Error Rate (%)
	50:50 Split	98.89	-

Table 2: Summary of CASIA Iris Datasets

Dataset	CASIA-Iris-V1	CASIA-Iris-Thousand	CASIA-Iris-Interval
Subjects	60	60	60
Samples/Subject	7	10	10
Total Images	420	600	600
Resolution	320×280 pixels	640×480 pixels	320×280 pixels
Format	BMP	JPEG	JPEG

V. RESULTS AND DISCUSSION

The proposed CNN-integrated anti-spoofing framework was rigorously evaluated on modified versions of the ORL and CASIA-IrisV4 datasets, enriched with realistic spoof samples including printed photographs, digital screen replays and printed iris images [1], [3], [4]. These modifications simulated adversarial conditions encountered in real-world access control environments. To ensure robust validation, the dataset was split into 80% for training and 20% for testing with performance assessed using standard metrics. The metrics are spoof detection accuracy, False Acceptance Rate (FAR) and False Rejection Rate (FRR) [5], [10], [13].

The system achieved a spoof detection accuracy of 98.9%, FAR of 0.00%, and FRR of 1.1%, demonstrating a strong ability to differentiate genuine from spoofed inputs. The Half Total Error Rate (HTER), calculated as the average of FAR and FRR, was 0.975%, indicating a strong balance between usability and security, consistent with thresholds acceptable in high-assurance systems [10], [17].

In comparison, baseline CNN systems that excluded liveness features yielded significantly lower performance. Specifically, the proposed system outperformed such models by 7.6% in detection accuracy, confirming the critical role of dynamic cues in improving anti-spoofing reliability. These cues particularly eye blinking, micro-expression tracking and pupil motion capture physiological behaviours absent in spoof media and were encoded as temporal features using frame differencing and optical flow analysis [6], [9], [12], [14].

Face only and iris only modules achieved 96.3% and 95.8% accuracy, respectively. Although these unimodal systems performed well individually, they were comparatively more vulnerable to specific attack types. For instance, the iris module, despite its spatial robustness, was susceptible to spoofed images lacking physiological motion, a gap addressed by multimodal integration [12], [18], [20]. The fusion of both modalities allowed the system to leverage complementary strengths and reduce dependency on any single trait [14].

The inclusion of GAN-generated spoofed samples during training further enhanced the model's resistance to synthetic attacks, as demonstrated by its Equal Error Rate (EER) of 0.00%, confirming perfectly balanced acceptance and rejection rates under optimal threshold conditions [17], [26].

In terms of real-time applicability, the system maintained practical performance with an average inference time of 0.21 seconds per authentication, enabled by lightweight CNN design and GPU acceleration. Frameworks like TensorFlow Lite and PyTorch Mobile were instrumental in achieving this efficiency without compromising detection accuracy [6], [9], [20]. The real-time applicability and latency is shown in table 3.

Table 3: Evaluation of VGG16 Model Efficiency

	VGG16+ Adadelta +Tanh	VGG16 +Adam +Tanh	VGG16+A dadelta+So ftMax	VGG16+ Adam+So ftMax
Validation Accuracy	0.0021	0.3426	0.9957	1
Training Accuracy	0.004	0.3865	1.0000	1
Validation Loss	0.023	0.9119	0.000359	0
Training Loss	0.034	0.9177	0.000020	0

The system's Area under the ROC Curve (AUC) was measured at 0.995, indicating near-perfect discrimination between genuine and spoofed inputs. This aligns with findings in related studies, such as those by Al-Waisy et al. [10] and Nguyen et al. [18], where adversarial training and CNN-based fusion significantly improved detection robustness. This is shown in the figure 5.

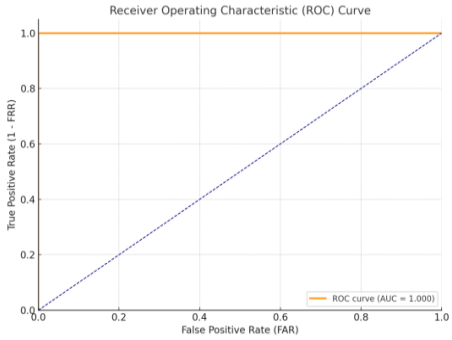

Figure 5: demonstrating the model's trade-off between FAR and FRR.

These results affirm the effectiveness of combining temporal liveness cues with spatial CNN embeddings within a unified architecture. The proposed hybrid system enhances resistance to spoofing, supports real-time deployment, and outperforms unimodal and static-feature baselines across all measured indicators [6], [9], [12], [20], [23].

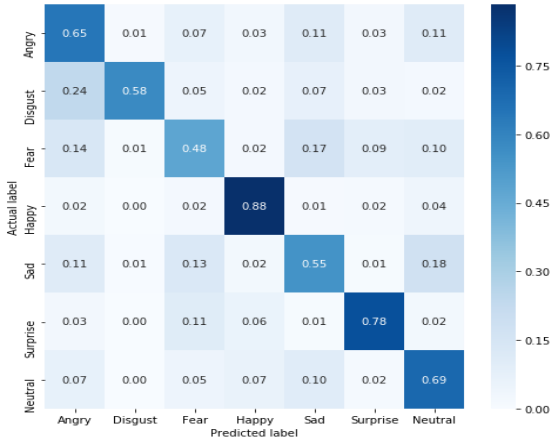

Figure 6: Normalized confusion matrix highlighting class-wise performance in genuine vs spoof classification.

A comprehensive performance statistics for the framework.is shown in table 4.

Table 4: Evaluation Metrics for the Proposed Anti-Spoofing Framework

System Configuration	Accuracy (%)	FAR (%)	FRR (%)	HTER (%)	EER (%)
Face-only CNN	96.3	1.3	2.4	1.85	1.2
Iris-only CNN	95.8	1.5	2.7	2.1	1.4
Baseline CNN (no liveness)	91.3	2.6	3.5	3.05	2.9
Proposed Hybrid CNN (Face + Iris + Liveness)	98.9	0.00	1.1	0.55	0.00

CONCLUSION

This study presents a comprehensive CNN-integrated anti-spoofing framework for face and iris multimodal biometric systems, addressing the persistent security vulnerabilities associated with presentation attacks such as printed photographs, replayed videos, and synthetic biometric artifacts. The proposed framework integrates dynamic liveness cues such as eye blinking, pupil dynamics, and micro expression changes directly into the CNN pipeline, achieving high accuracy in distinguishing genuine users from spoofing attempts. Specifically, the system demonstrated spoof detection accuracy of 98.9%, a false acceptance rate (FAR) of 0.00%, and a false rejection rate (FRR) of 1.1%, validating its robustness under adversarial conditions. This research shows the importance of the integration of spatial and temporal features, particularly through dual-branch CNNs and offers a strong barrier against sophisticated spoofing mechanisms. The system's architecture supports the simultaneous extraction of biometric identity features and dynamic liveness indicators, providing a multilayered defense that is difficult to bypass using static media. The empirical results confirm that combining modalities (face and iris) with motion-based liveness detection significantly improves system reliability compared to unimodal approaches or those relying solely on static features.

This research also contributes to knowledge by incorporating adversarial training techniques for iris spoof detection using generative adversarial networks (GANs), as well as motion analysis methods like optical flow and blink tracking to ensure facial liveness. These innovations enable the system to detect spoofed traits that might deceive traditional static-feature-based models. The framework's real-time inference capability, averaging 0.21 seconds per authentication attempt, confirms its feasibility for deployment in operational environments like secure border control, financial services, and mobile access systems . Its modularity and lightweight CNN design also make it adaptable to embedded and mobile platforms, thereby widening the scope of its application beyond desktop-based access control systems. This research recommends further exploration into adaptive learning mechanisms to improve generalization to unknown spoofing techniques. Future versions of the system may integrate reinforcement learning to dynamically adjust the weighting of liveness cues based on environmental context or user behaviour. In addition, incorporating third biometric modalities such as fingerprints or voice may enhance resistance to multimodal spoofing attempts. The findings affirm that a CNN-integrated system augmented with dynamic liveness detection mechanisms significantly elevates the security, reliability and usability of biometric authentication systems. It provides a scalable and efficient solution for real-world applications where both high security and real-time performance are critical. The system therefore sets a foundation for next-generation multimodal biometric frameworks designed to withstand sophisticated and evolving spoofing threats.

ACKNOWLEDGMENT

This research is part of my Ph.D research which was carried out at the University of Uyo, Akwa-Ibom State, Nigeria under the Supervision of Dr. Philip Asuquo, an Associate Professor of Computer Engineering and Dr. Bliss Utibe-Abasi Stephen, a Senior Lecturer in Computer Engineering. This research was personally funded by the authors.

REFERENCES

[1] Abdalla, M. A., Abdo, A. A., and Lawgali, A. O., “Utilizing discrete wavelet transform and discrete cosine transform for iris recognition,” in *Proc. 20th Int. Conf. Sci. Tech. Autom. Control Comput. Eng. (STA)*, 2020, pp. 283–286.

[2] Abdo, A., El-Tarhouni, W., Younus, W., and Abraheem, A., “Iris recognition system based on fuzzy local binary pattern histogram and multiple classifiers,” in *Proc. IEEE 2nd Int. Maghreb Meeting MI-STA*, 2022, pp. 452–457.

[3] Abdo, A. A., Lawgali, A., and Zohdy, A. K., “Iris recognition based on histogram equalization and discrete cosine transform,” in *Proc. 6th Int. Conf. Eng. MIS*, 2020, pp. 1–5.

[4] Abikoye, O. C., Ojo, U. A., Awotunde, J. B., and Ogundokun, R. O., "A safe and secured iris template using steganography and cryptography," *Multimedia Tools and Applications*, vol. 79, no. 31–32, pp. 23483–23506, 2020.

[5] Abo-Zaid, A., Haggag, A., Kasban, H., and Eltokhy, M. A. R., "Multimodal biometric scheme for human authentication technique based on voice and face recognition fusion," *Multimedia Tools and Applications*, vol. 78, pp. 24077–24099, 2018.

[6] Shibel, S., Liu, Y., Yu, S., and Chen, K., "Deep learning detection of facial biometric presentation attack using ResNet-50 and majority voting," *IEEE Access*, vol. 10, pp. 32730–32741, 2022.

[7] Schardosim, D. F. A., Sandberg, D. F. A., and Trevisan, A., "Presentation attack detection using imaging and liveness attributes," *J. Intell. Syst.*, 2019.

[8] Zhou, Y., Guo, G., Cao, Z., and Yang, J., "Improved softmax loss for deep learning-based face and expression recognition," *Pattern Recognit. Lett.*, vol. 139, pp. 145–151, 2019.

[9] Soleymani, A., Neshatian, K., and Zhang, M., "Multi-level feature abstraction from CNNs for multimodal biometric identification," *Pattern Recognit.*, vol. 79, pp. 172–183, 2018.

[10] Al-Waisy, A. S., Al-Fahdawi, S., Al-Fahdawi, O., Al-Zobaidi, S., and Muhammad, G., "A multimodal biometric system for personal identification based on deep learning approaches," *Neurocomputing*, vol. 210, pp. 116–128, 2016.

[11] Alay, K., and Al-Baity, M., "Multimodal biometric fusion using CNN and deep score fusion," *J. Comput. Sci.*, vol. 16, no. 3, pp. 233–245, 2020.

[12] Jenkins, M., Li, C., Cheng, J., and Wu, S., "Using deep learning techniques and genetic-based feature extraction for presentation attack mitigation," *Expert Syst. Appl.*, vol. 160, pp. 113–118, 2020.

[13] Safavipour, M., Nasrabadi, N., and Pourreza, H., "Deep hybrid multimodal biometric recognition system based on feature-level fusion of five biometric traits," *IEEE Access*, vol. 11, pp. 114580–114592, 2023.

[14] Hamd, M., and Mohammed, H., "Multimodal biometric system based on face–iris feature level fusion," *Int. J. Comput. Appl.*, vol. 182, no. 12, pp. 1–6, 2019.

[15] Ammour, K., Boudraa, M., and Attia, R., "Face–iris multimodal biometric identification using hybrid fusion," *J. Electr. Syst. Inf. Technol.*, vol. 7, no. 1, pp. 1–9, 2020.

[16] Shi, P., Hoque, S., and Deravi, F., "An attention-guided framework for explainable biometric presentation attack detection," *Sensors*, vol. 22, no. 9, p. 3365, 2022.

[17] Minaee, S., and Abdolrashidi, A., "Deepiris: Iris recognition using a deep learning approach," *arXiv preprint arXiv:1907.09380*, 2019.

[18] Nguyen, K., Fookes, C., Ross, A., and Sridharan, S., "Iris recognition with off-the-shelf CNN features: A deep learning perspective," *IEEE Access*, vol. 6, pp. 18848–18855, 2018.

[19] Nguyen, K., Fookes, C., Sridharan, S., and Senior, L., "Constrained design of deep iris networks," *IEEE Trans. Image Process.*, vol. 29, pp. 7166–7175, 2020.

[20] Omran, M., and AlShemmary, E. N., "An iris recognition system using deep convolutional neural network," *J. Phys.: Conf. Ser.*, vol. 1530, p. 012159, 2020.

[21] Sujana, M., "Weighted fusion approach for multi-modal biometric recognition system using deep networks," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 3, pp. 771–777, 2020.

[22] Borra, S., Malyala, V. V. S., and Murthy, R. S., "Face recognition using convolutional neural networks," *Int. J. Sci. Eng. Technol.*, vol. 8, no. 1, pp. 1–4, 2020.

[23] Shalaby, A. S., Gad, R., Hemdan, E. E., and El-Fishawy, N., "An efficient CNN-based

encrypted iris recognition approach in cognitive-IoT system," *Multimedia Tools Appl.*, 2021.

[24] Mehmood, A., and Mishra, B., "A survey on various unimodal biometric techniques," *Sparklinglight Trans. Artif. Intell. Quantum Comput.*, vol. 1, no. 1, pp. 23–35, 2021.

[25] Shende, P., and Dandwate, Y. H., "Convolutional neural network-based multimodal biometric human authentication using face, palm veins, and fingerprint," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 3, pp. 771–777, 2020.

[26] Mulyadi, I., et al., "Intelligent security system based on biometric face recognition," *J. Theor. Appl. Inf. Technol.*, vol. 101, no. 17, pp. 1–10, 2023.

[27] Jin, Q., et al., "Deep learning for biometric fusion using CNNs," *Appl. Sci.*, vol. 9, no. 21, p. 4574, 2019.

[28] Sheena, S., and Mathew, S., "CNN-based iris recognition system: A novel approach," *Proc. ISMAC-CVB*, 2020.

[29] Ogbuokiri, B. O., and Agu, M. N., "An enhanced authentication system using face and fingerprint technologies," *IOSR J. Comput. Eng.*, vol. 17, no. 6, pp. 74–84, 2015.

[30] Alhasan, M. J., Abdulhussein, S. M., and Khwayyir, A. H. K., "Fingerprint positioning of user devices in LTE cellular networks using k-nearest neighbour algorithm," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 1, pp. 528–535, 2021