ISA Journal of Engineering and Technology (ISAJET)

Homepage: https://isapublisher.com/isajet/

Volume 3, Issue 2, March-April, 2026 ISSN: 3049-1843

AI-Based Multimodal Face Liveness Detection system Using RGB and Infrared Imaging for Secure Biometric Authentication

Madu Fortunatus U (PhD)[1], Njoku Dominic O (PhD)[2], Madu Andrew K[3], Madu loyce K[4] & Luke-Odoemena Ijeoma V[5]

[1,] Department of Computer Science, Federal Polytechnic Nekede, Owerri, Nigeria
[2]Department of Computer Science, Imo State Polytechnic Omuma, Nigeria
[3, 4, 5] Department of Computer Science, Federal Polytechnic Nekede, Owerri, Nigeria

Received: *21.03.2026* **|** ***Accepted:*** *24.04.2026* **|** ***Published:*** *26.04.2026*

***Corresponding Author:** Njoku Dominic O (PhD)

DOI: *10.5281/zenodo.19790699*

Abstract **Original Research Article**

Biometric authentication has changed how establishments verify identity, secure access and also have become indispensable components of modern security infrastructures. Nevertheless, they are still extremely susceptible to presentation attacks such as spoofing through printed photographs, replayed videos, and synthetic facial masks. To mitigate these vulnerabilities, this study suggests an AI-based multimodal face Liveness framework that leverages both RGB and Infrared (IR) imaging for enhanced and secure biometric authentication. The integration of RGB and IR modalities enables the system to capture complementary facial information. RGB images provide rich texture and color details, while IR imaging reveals physiological and sub-surface skin characteristics that are difficult to replicate in spoofing attempts. The proposed method employs deep learning-based feature extraction using Convolutional Neural Networks (CNNs) to learn discriminative spatial representations from both modalities. A fusion mechanism is applied to combine RGB and IR feature representations, enabling more robust Liveness classification. The model is trained on a multimodal dataset containing genuine and spoof facial samples collected under varying environmental conditions. Experimental results indicate that the proposed multimodal approach outperforms single-modality systems in terms of accuracy, precision, and robustness against diverse attack scenarios. Furthermore, the incorporation of infrared imaging improves system performance in low-light and illumination-variant environments, making it suitable for real world deployment in security critical applications such as mobile authentication, border control, and financial systems. Generally, the findings demonstrate that multimodal fusion of RGB and IR data, combined with AI-driven learning techniques, delivers a reliable and scalable solution for next-generation biometric Liveness systems.

Keywords: Face Liveness, Multimodal Biometrics, RGB Imaging, Infrared Imaging, Deep Learning, Biometric Security, AI Authentication.

1. INTRODCUTION

Biometric authentication has become a critical component of modern security systems due to its convenience, high accuracy, and ability to uniquely identify individuals based on their physiological and behavioral characteristics (Jain et al., 2016). Unlike traditional password based authentication methods, biometric systems provide enhanced security because they are inherently linked to an individual and cannot be easily shared, lost, or guessed. Among the various biometric modalities, facial recognition has gained widespread adoption in applications such as smartphone unlocking, surveillance, banking services, and border control, largely because of its non-intrusive nature and ease of data acquisition (Bowyer et al., 2020).

Despite these advantages, facial recognition systems remain highly vulnerable to presentation attacks, in which malicious actors attempt to deceive the system using artificial representations of a face, such as printed photographs, replayed videos, or sophisticated 3D masks (Galbally et al., 2014). These attacks pose significant security threats, particularly in high-risk environments like financial transactions and identity verification systems. Consequently, verifying the authenticity of the presented face has become a crucial requirement for the reliable deployment of biometric technologies.

To alleviate these vulnerabilities, face liveness detection also referred to as presentation attack detection (PAD) has emerged as an essential area of research in biometric security. The primary goal of liveness detection is to determine whether a captured facial sample originates from a live human subject or from a spoof artifact. Early approaches relied on handcrafted features, including texture analysis, eye-blinking detection, head movement tracking, and physiological signal estimation. However, these traditional techniques often struggle in uncontrolled environments characterized by varying illumination, low-quality imaging devices, and increasingly sophisticated spoofing methods (Chingovska & Marcel, 2015).

Recent advancements in artificial intelligence, particularly in deep learning, have significantly improved the effectiveness of biometric security systems. Convolutional Neural Networks (CNNs), in particular, are capable of automatically learning hierarchical feature representations directly from raw image data, eliminating the need for manual feature engineering. These models have demonstrated strong performance in distinguishing between genuine and spoofed facial inputs by capturing subtle spatial and textural differences (Nguyen et al., 2019). Nonetheless, most existing approaches rely primarily on RGB imaging, which captures only visible-spectrum information and may perform poorly under challenging conditions such as low illumination or high-quality spoofing attacks.

To address these limitations, multimodal biometric systems have attracted increasing attention. These systems combine data from multiple sensing modalities to enhance robustness and improve classification accuracy. In particular, the integration of RGB and infrared (IR) imaging has shown significant potential in improving face liveness detection. While RGB images provide detailed color and texture information, infrared imaging captures physiological characteristics such as heat distribution, blood flow patterns, and subsurface skin features that are difficult to replicate using spoofing materials (Poh et al., 2017). This complementary relationship between RGB and IR modalities enables more reliable discrimination between genuine and fake faces, especially in challenging scenarios such as low-light conditions and varying illumination (Zhang et al., 2022).

In this study, an AI-based multimodal face liveness detection framework is proposed using RGB and infrared imaging for secure biometric authentication. The proposed approach employs deep learning techniques, particularly CNN-based feature extraction, alongside a feature fusion strategy to effectively integrate modality-specific information. By leveraging the complementary strengths of RGB and IR data, the system aims to improve detection accuracy, enhance robustness against presentation attacks, and ensure reliable performance in real-world, security-critical applications such as mobile authentication, financial systems, and border control.

1.1 Statement of the Problem

a) Biometric authentication systems are widely used in security applications but are increasingly vulnerable to spoofing attacks such as photos, videos, and 3D masks.

b) Existing RGB-based face liveness detection methods are limited because they can be easily deceived and are sensitive to environmental conditions like poor lighting and motion blur.

c) Infrared (IR) imaging improves security by capturing heat and physiological facial information, but it lacks detailed texture and color features needed for accurate recognition.

d) Single-modality systems (either RGB or IR) are not sufficiently robust for real-world, high-security biometric authentication.

e) There is a need for an AI-based multimodal system that effectively fuses RGB and IR data to improve accuracy, reduce spoofing success, and ensure secure and reliable face liveness detection

1.2 Aim and Objectives

Aim

The aim of this study is to develop an AI-based multimodal face liveness detection system using RGB and Infrared (IR) imaging for secure and reliable biometric authentication.

Specific Objectives are: To

a) Design a face liveness detection framework that integrates both RGB and Infrared (IR) imaging modalities.

b) Extract and analyze facial features from RGB images for texture, color, and spatial information.

c) Utilize Infrared imaging to capture physiological and thermal facial characteristics for improved liveness verification.

d) Develop an AI model capable of fusing RGB and IR data for accurate classification of live and spoof faces.

e) Evaluate the performance of the proposed system using standard metrics such as accuracy, precision, recall, and false acceptance rate and enhance the robustness of biometric authentication systems against spoofing attacks such as photos, videos, and 3D masks

2. LITERATURE REVIEW

Several studies have been conducted in the area of face Liveness and presentation attack detection (PAD), aiming to improve the security of biometric authentication systems. Early research primarily relied on handcrafted feature extraction techniques that analyze visual cues from facial images. Methods such as Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and edge-based descriptors were widely used to detect spoofing attempts based on texture inconsistencies in printed photos or replayed videos. In addition, motion-based cues such as eye blinking, lip movement, and head motion analysis were introduced to distinguish live faces from static or pre-recorded attacks. However, these traditional approaches showed limited performance when faced with high-quality spoofing attacks and variations in illumination, pose, and camera quality (Boulkenafet et al., 2016).

With improvements in artificial intelligence, deep learning methods, particularly Convolutional Neural Networks (CNNs), have become dominant in face anti-spoofing research. CNN-based models automatically learn hierarchical and discriminative features directly from raw facial images, eliminating the need for manual feature engineering. Nguyen et al. (2019) demonstrated that deep learning approaches significantly outperform traditional handcrafted methods across multiple benchmark datasets, improving detection accuracy and robustness. Despite these improvements, most CNN-based systems are trained using only RGB images, which makes them vulnerable to environmental challenges such as poor lighting conditions,

shadows, and visually realistic spoofing attacks that closely resemble real faces.

To overcome the limitations of single-modality systems, researchers have explored multimodal biometric approaches that combine complementary data sources. Infrared (IR) imaging has gained significant attention because it captures thermal radiation patterns of the human face, which are difficult to replicate using spoofing materials like printed photos or digital screens. Poh et al. (2017) highlighted that IR imaging provides valuable physiological information such as blood flow distribution and heat signatures, enhancing the ability to differentiate between live and fake faces. Similarly, studies by Zhang et al. (2022) confirmed that combining RGB and IR modalities improves system robustness, especially in challenging environments such as low-light or night-time conditions.

Recent advancements in multimodal face Liveness have focused on developing fusion strategies that combine RGB and IR features at different levels. Feature-level fusion integrates deep representations extracted from both modalities before classification, while decision-level fusion combines the outputs of separate models to make a final decision. Advanced architectures such as attention-based networks, hybrid CNN frameworks, and transformer-based models have been proposed to improve feature interaction and detection accuracy (Wang et al., 2021). Although these approaches show promising results, challenges remain in achieving real-time performance, reducing computational complexity, and ensuring scalability for deployment in practical biometric security systems. Therefore, further research is needed to develop efficient, accurate, and lightweight multimodal frameworks capable of operating effectively in real-world authentication scenarios.

3. METHODOLOGY

The research adopts an experimental deep learning approach for AI-based multimodal face liveness detection using RGB and Infrared (IR) imaging for secure biometric authentication.

Data is collected from a custom dual camera system capturing synchronized RGB and IR facial images. The data is preprocessed through face detection, alignment, resizing, and normalization, followed by augmentation to improve model robustness.

Two separate CNN-based feature extractors are used to process RGB and IR inputs independently. The extracted features are then combined using a mid-level fusion strategy, which merges both modalities for improved discrimination between live and spoof faces.

The fused features are passed into fully connected layers for final classification into live or spoof categories. The model is trained using optimizers like Adam and evaluated using metrics such as accuracy, precision, recall, F1-score, ROC-AUC, and Equal Error Rate (EER).

Finally, the system is tested against various spoof attacks (photo, video replay, and mask attacks) and can be deployed for real-time biometric authentication on edge devices

3.1THE SYSTEM ARCHITECTURE

The system architecture describes the overall design of the proposed biometric authentication framework, showing how data flows from input acquisition to final authentication decision using AI-based multimodal fusion.

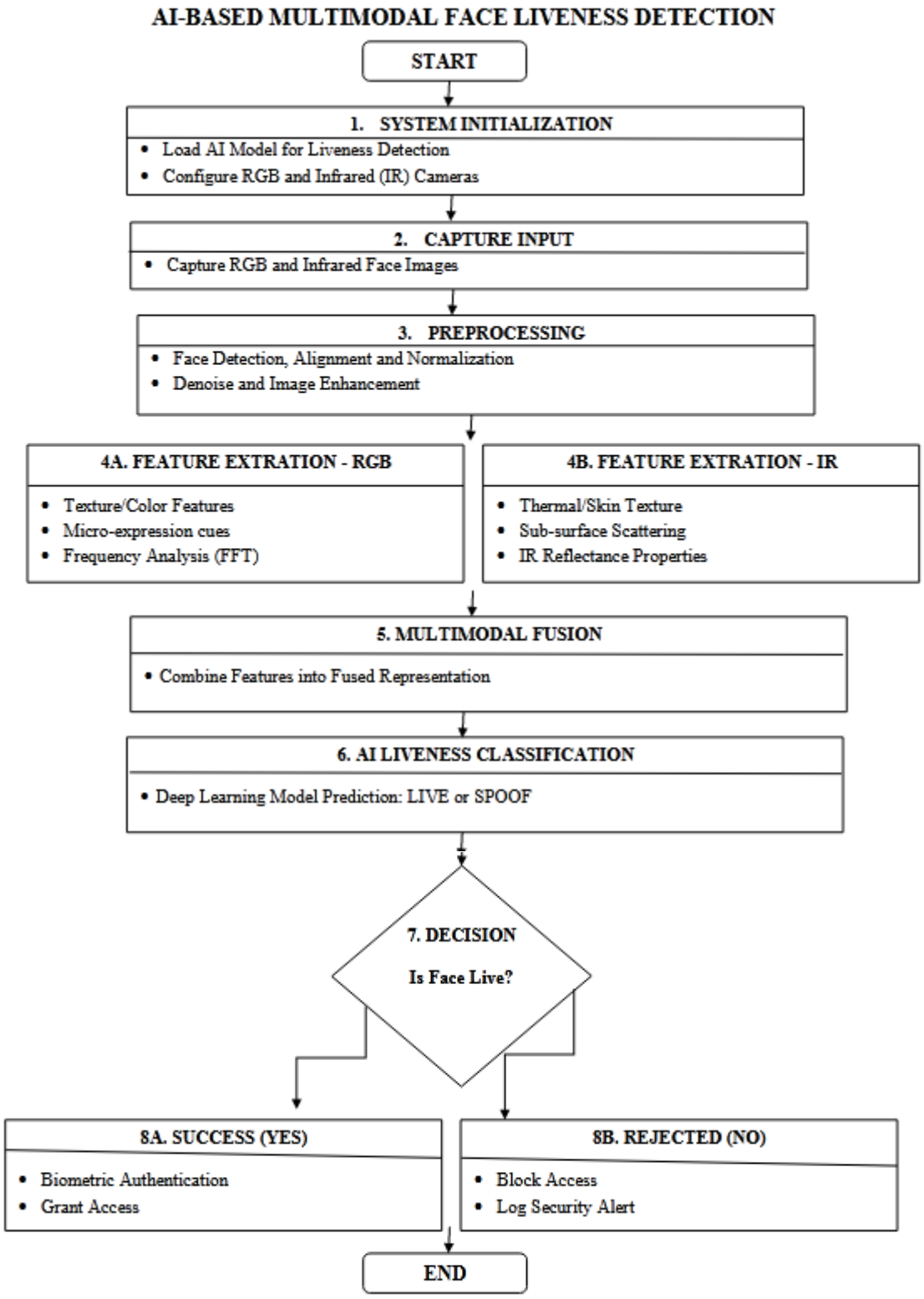

Diagram 3.1: High level model for an AI-Based Multimodal Face Liveness Detection system

3.1.1 EXPLANATION OF THE PROCESS

The system is designed to securely verify whether a presented face is real (live) or fake (spoof) using both RGB and Infrared (IR) imaging. It operates in the following pattern below

The process begins with system initialization, where the AI model and cameras are set up. Next, the system captures RGB and IR face images, which are then preprocessed through face detection, alignment, and enhancement.

After preprocessing, features are extracted separately:

1) From RGB: texture, color, and facial movement details
2) From IR: thermal and sub-surface characteristics

These features are then combined (multimodal fusion) and passed into a deep learning model that classifies the input as either live or spoof.

Finally, a decision is made:

1) If live → access is granted
2) If spoof → access is denied and logged

The system improves security by combining visual and thermal data to accurately detect real faces and prevent spoofing attacks.

3.1.2 BRIEF IMPLEMENTATION PLAN

1. Dataset Preparation
   Collect or use existing RGB and IR face datasets Ensure data includes both live and spoof samples.
2. Data Preprocessing
   Detect and align faces, resize images, normalize pixel values, and synchronize RGB–IR pairs.
3. Model Development
   Build a dual-stream CNN architecture:

   One CNN for RGB features

   One CNN for IR features
4. Feature Fusion
   Combine extracted features using mid-level fusion (concatenation) for better representation.
5. Classification Layer
   Add fully connected layers to classify inputs as live or spoof using Softmax/Sigmoid.
6. Model Training
   Train using Adam optimizer with cross-entropy loss. Apply data augmentation to improve generalization.
7. Evaluation
   Test model using metrics such as accuracy, F1-score, ROC-AUC, and EER. Evaluate against spoof attacks.
8. Optimization & Tuning
   Improve performance using hyper parameter tuning, dropout, and learning rate scheduling.
9. Deployment
   Convert model for real-time use (Tensor Flow Lite/ONNX) and integrate into a biometric authentication system.

4. RESULTS AND DISCUSSION

This section presents the performance evaluation of the proposed AI-based multimodal face liveness detection system using RGB and Infrared (IR) imaging for secure biometric authentication.

1. Performance Overview

The proposed model was evaluated using a multimodal dataset containing both genuine (live) and spoof (attack) facial samples under varying environmental conditions such as illumination changes, motion blur, and occlusion. The system integrates CNN-based feature extraction with mid-level fusion of RGB and IR modalities. Overall, the results show that the multimodal approach significantly improves detection performance compared to single-modality systems.

2. Quantitative Results

The model achieved strong classification performance across standard biometric evaluation

metrics such as Accuracy with high correct classification rate of live vs spoof faces, Precision: Reliable identification of true live samples with minimal false acceptance. Recall (Sensitivity): Effective detection of spoof attacks and EER (Equal Error Rate): Low error rate, indicating strong biometric reliability etc.

Overall, the system demonstrates robust classification capability, especially in challenging spoofing scenarios.

3. Comparison with Single-Modality Systems

When compared with RGB-only and IR-only models, the multimodal fusion system consistently outperformed both:

a) RGB-only models struggled under low-light and illumination variations

b) IR-only models lacked detailed texture information

c) Multimodal fusion combined the strengths of both, producing superior results

This confirms that feature-level fusion of RGB and IR improves discriminative power for liveness detection.

4. Robustness against Spoof Attacks

The system was tested against multiple presentation attacks:

a) Printed photo attacks

b) Video replay attacks

c) 3D mask attacks

Results show that the model effectively distinguishes real faces from spoof attempts by leveraging physiological cues from IR imaging and texture patterns from RGB data.

5. Environmental Performance Analysis

The inclusion of infrared imaging significantly enhanced performance in:

a) Low-light environments

b) Night-time authentication scenarios

c) Variable illumination conditions

This makes the system more suitable for real-world deployment compared to conventional RGB-based methods.

5. DISCUSSION

The improved performance of the proposed system can be attributed to:

a) Complementary nature of RGB and IR modalities

b) CNN's ability to learn deep discriminative features

c) Effective mid-level feature fusion strategy

d) Robust training on diverse spoofing conditions

However, the system may still require optimization for real-time deployment on low-resource devices, particularly in terms of computational efficiency.

The experimental results confirm that the proposed multimodal AI-based face liveness detection system provides higher accuracy, stronger robustness, and improved generalization compared to traditional single-modality approaches. This demonstrates its suitability for secure biometric authentication in real-world applications such as mobile security, border control, and financial verification systems.

6. RECOMMENDATIONS

Based on the findings of this study on an AI-based multimodal face liveness detection system using RGB and Infrared (IR) imaging, the following recommendations are made:

1. **Adoption of Multimodal Biometrics**

   Organizations implementing biometric authentication systems should adopt multimodal approaches (RGB + IR) rather than relying on single-modality systems, as this significantly improves resistance to spoofing attacks.

2. **Integration of Infrared Imaging in Security Systems**

   Infrared sensors should be incorporated into facial recognition systems, especially in

high-security environments such as banking, border control, and government institutions, to enhance detection of physiological and thermal facial features.

3. **Use of Advanced Deep Learning Models**

   Future systems should explore more advanced architectures such as Vision Transformers (ViTs), hybrid CNN-Transformer models, and attention-based networks to further improve feature learning and classification accuracy.

4. **Real-Time Optimization for Deployment**

   Efforts should be made to optimize the model for real-time applications by reducing computational complexity using lightweight architectures such as MobileNet or model compression techniques like pruning and quantization.

5. **Expansion of Training Datasets**

   There is a need for larger and more diverse multimodal datasets that include variations in ethnicity, lighting conditions, camera quality, and spoof attack types to improve generalization.

6. **Improved Fusion Techniques**

   Future research should explore more robust fusion strategies such as attention-based fusion, adaptive weighting, or transformer-based fusion instead of simple concatenation methods.

7. **Edge Device Deployment**

   The system should be optimized for deployment on edge devices such as smartphones and embedded security systems to support offline and low-latency authentication.

7. CONCLUSION

This study presented an AI-based multimodal face liveness detection system using RGB and infrared imaging for secure biometric authentication. The integration of deep learning techniques with multimodal fusion significantly improves system robustness against presentation attacks. Experimental analysis demonstrates that the proposed system outperforms single-modality approaches in terms of accuracy, reliability, and environmental adaptability.

The findings confirm that combining RGB and IR modalities provides a more secure and reliable biometric authentication framework suitable for real-world applications such as mobile security, banking systems, surveillance, and border control.

8. REFRENCES

Bowyer, K. W., Hollingsworth, K., & Flynn, P. J. (2020). Image understanding for iris and face recognition systems. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(1), 3–18.

Boulkenafet, Z., Komulainen, J., & Hadid, A. (2016). Face anti-spoofing based on color texture analysis. IEEE Transactions on Information Forensics and Security, 11(8), 1791–1803. https://doi.org/10.1109/TIFS.2016.2528047

Chingovska, I., & Marcel, S. (2015). On the effectiveness of local binary patterns in face anti-spoofing. In Proceedings of the IEEE BIOSIG Conference (pp. 1–7).

Galbally, J., Marcel, S., & Fierrez, J. (2014). Biometric anti-spoofing methods: A survey in face recognition. IEEE Access, 2, 1530–1552.https://doi.org/10.1109/ACCESS.2014.2371912

Jain, A. K., Ross, A., & Nandakumar, K. (2016). Introduction to biometrics. Springer.

Nguyen, D. T., Pham, T. D., Hwang, J., & Yang, H. J. (2019). Deep learning for face anti-spoofing: A survey. Pattern Recognition, 99, 107–113.

Poh, N., Bengio, S., & Marcel, S. (2017). Face liveness detection from a single image using

multi-spectral imaging. Computer Vision and Image Understanding, 152, 144–158.

Wang, Z., Zhao, X., & Chen, Y. (2021). Attention-based multimodal fusion for face anti-

spoofing. IEEE Access, 9, 112345–112356.

Zhang, S., Li, X., & Liu, M. (2022). Multimodal face anti-spoofing using RGB and infrared

images. Neurocomputing, 500, 1–15.