A Deep Learning-Driven Multimodal Biometric Medical Image Protection System with Secure Encryption

S. N. Kavitha
Department of Information Science and Engineering, RV College of Engineering, Bengaluru, India
kavithasn@rvce.edu.in (corresponding author)

K. Vanishree
Department of Information Science and Engineering, RV College of Engineering, Bengaluru, India
vanishreek@rvce.edu.in

Received: 8 November 2025 | Revised: 12 January 2026 and 14 February 2026 | Accepted: 15 February 2026

 DOI: https://doi.org/10.48084/etasr.16114

ABSTRACT

The Industrial Internet of Things (IIoT) has transformed healthcare by enabling remote diagnosis and efficient exchange of medical data, but it also raises serious privacy and security concerns. Encryption alone cannot fully protect medical images after decryption, while unimodal biometrics security lacks reliability. To overcome these issues, this work proposes BioMedShield, a deep learning-driven multimodal biometric medical image protection system that integrates encryption, data hiding, and biometric fusion for comprehensive security. The framework uses Dual-SegNet (DS-Net) for accurate medical image segmentation and MSE-Net for robust biometric feature extraction. DS-Net achieves high performance with 98.8% accuracy and 98.7% recall. Secure Biometric Blowfish Encryption (SBBE) further ensures strong confidentiality and resistance to unauthorized access in IIoT healthcare systems.

Keywords-deep learning; multimodal biometrics; medical image protection; data hiding; encryption; Industrial Internet of Things (IIoT); secure healthcare system

I. INTRODUCTION

Recent advances in Artificial Intelligence (AI) have greatly improved healthcare, especially in deep learning-based medical image analysis [1]. Traditional encryption and access control methods are often not sufficient to protect complex multimodal medical data [2]. To overcome this, the combination of multimodal biometric authentication with advanced encryption techniques offers a more reliable solution to maintain patient confidentiality and diagnostic accuracy [3]. Deep learning-driven systems can effectively extract and fuse biometric features from multiple sources, such as facial traits, retinal patterns, and fingerprint-like medical features, using convolutional and transformer-based models [4]. This multimodal strategy improves authentication accuracy and reduces false acceptance and rejection rates.

Linking access to medical images with biometric verification ensures that only authorized doctors or patients can decrypt sensitive data, strengthening identity-based security [5]. In addition, secure encryption methods such as chaotic mapping, deep feature hashing, and homomorphic encryption help preserve data integrity during storage and transmission [6]. Adaptive, deep learning-based key generation further improves protection [7].

In general, integrating deep learning, multimodal biometrics, and secure encryption provides end-to-end medical image security, supporting safe and efficient data access in IIoT-enabled healthcare environments [8-12]. The proposed BioMedShield system addresses the weaknesses of traditional security models by combining multimodal biometrics, deep feature learning, and hybrid encryption. This approach enhances privacy, image ownership verification, and defense against cyber-attacks, thus creating a reliable and secure digital healthcare environment. The key major contributions of this study are:

- BioMedShield verifies users with OTP, preventing any unauthorized system access.
- The framework consists of segmentation, feature extraction, data hiding, and encryption stages.
- Dual-SegNet segments lesions using CalibNeXt encoder and Residual Swin Transformer.
- Lesions are hidden in non-lesion areas using the LSB method securely.
- SBBE encrypts medical images using biometric keys and device information.

- Ensures strong privacy, integrity, and security for medical data.

In [13], a deep learning-based warning medical imaging system was based on multimodal biometrics, data hiding, and encryption. DeepENC [14] is a deep learning encryption model based on multimodal features for transmitting medical images securely. In [15], a Privacy-Preserving and Authenticating Framework of Biometric-based Systems (PPAF-BS) was based on Hybrid Deep Learning (HDL) with palmprint, ear, and face biometrics. In [16], a multimodal medical image protection system combined the data encryption and fusion procedures. Previously, a Fragmented Medical Health Record Security System [17] added encrypted patient data in a series of medical images through a modified LSB process. In [18], a Multi-Modal Biometric Authentication Model combined CNN and RNN networks with shared and modality-specific layers. In [19], a multimodal biometric system, based on DNN-ChOA, introduced palm and knuckle vein recognition with optimization of features. In [20], a secure multimodal biometric key generation system combined a CNN with bat optimization, called Separately Extracted Feature Fusion (SEFF-CNN-BO). Table I shows that where multimodal biometrics and deep learning improve security, efficient, scalable end-to-end solutions remain limited.

A. Research Gap

Despite significant advances in deep learning, multimodal biometrics, and medical image encryption, previous studies exhibit several critical limitations that necessitate further research. Most prior works focus on isolated components, such as biometric authentication, encryption, or data hiding, rather than providing a fully integrated end-to-end security framework. In addition, unimodal or limited multimodal biometric systems often suffer from reduced reliability, higher false acceptance/rejection rates, and vulnerability to spoofing attacks. Many encryption techniques ensure data security during transmission but fail to protect medical images after decryption, leaving sensitive information exposed. Furthermore, existing models lack scalability and adaptability for real-time deployment in Industrial Internet of Things (IIoT)-based healthcare environments. There is also limited emphasis on combining accurate medical image segmentation with secure data embedding and biometric-driven key generation within a unified architecture. Hence, a robust, scalable, and fully integrated system that simultaneously addresses segmentation accuracy, multimodal biometric fusion, secure data hiding, and strong encryption remains an open research challenge, which this study aims to resolve through the proposed BioMedShield framework.

TABLE I. RELATED STUDIES

Study - Year	Proposed model	Dataset type / Modality	Features	Performance metrics
[13] - 2024	Secure Healthcare Imaging System	Face & Iris Biometric, CT Images	Deep learning segmentation, LSB data hiding, multimodal biometrics, 2D chaotic encryption	High segmentation & encryption accuracy, improved robustness
[14] - 2024	DeepENC	Fingerprint & Iris, ROI-based CT Images	ROI-based encryption using UNet3+, multimodal fusion, 2D hybrid chaotic key generation	High encryption accuracy, key sensitivity, low time complexity
[15] - 2025	PPAF-BS Framework	Palmprint, Ear & Face Images	Hybrid Deep Learning (HDL), DCT & Lagrange interpolation for privacy	Accuracy = 96.4%, reduced database size, high privacy
[16] - 2025	Coupled Chaotic Mapping-Based Security Scheme	CT Images with EMR data	Logistic-Cubic chaotic map, image fusion+encryption, blind watermarking	NPCR = 99.61%, UACI = 33.46%, strong noise robustness
[17] - 2024	Fragmented Medical Health Record Protection	Multimodal medical images	Encrypted medical record embedding via modified LSB, multimodal image fragments	High PSNR, low MSE, strong SSIM, diagnostic integrity maintained
[18] - 2025	Multimodal Biometric Authentication Model	Face, Voice, Signature	CNN + RNN hybrid, shared & modality-specific layers, PCA fusion, GBM classification	High authentication accuracy, strong robustness
[19] - 2025	DNN-ChOA Multimodal System	Palm & Knuckle Vein images	GLCM + DWT features, Chimp optimization algorithm	Accuracy = 99.85%, Sensitivity = 98.25%, Specificity = 97%
[20] - 2024	SEFF-CNN-BO System	Iris, Face & Fingerprint	CNN with Bat Optimization, Attribute-based encryption	Accuracy = 99.56%, Recall = 99.64%, strong attack resistance

II. PROPOSED MODEL

Figure 1 shows the overall architecture of BioMedShield. Its step-by-step workflow is as follows:

- Step 1 - User Authentication and Access Control: Authorized users (doctors/patients) first verify their identity using OTP-based login, ensuring that only legitimate users can initiate the process and access medical data.
- Step 2 - Multimodal Biometric Acquisition and Key Generation: After login, biometric inputs are collected and processed using MSE-Net to extract robust features. These features are fused to generate a secure biometric key, which is later used in encryption.
- Step 3 - Medical Image Segmentation: The input CT image is processed using DS-Net to accurately segment lesion and non-lesion regions. This step is essential to identify safe regions for secure data embedding.
- Step 4 - Secure Data Hiding: Sensitive medical information (e.g., patient data or lesion details) is embedded into non-lesion regions using LSB along with timestamp encoding, ensuring no diagnostic information is altered.
- Step 5 - Biometric-Based Image Encryption: The segmented and data-embedded image is encrypted using SBBE. The encryption key is derived from multimodal biometric features and device-specific parameters, ensuring strong confidentiality.
- Step 6 - Secure Storage and Transmission: The encrypted image is safely stored or transmitted, preventing unauthorized interception or tampering.

- Step 7 - Decryption and Authorized Access: At the receiver end, only authenticated users with matching biometric inputs can regenerate the key, decrypt the image, and retrieve the hidden data securely.
- Step 8 - Performance Evaluation: The system is evaluated using segmentation metrics (accuracy, recall), security metrics (NPCR, UACI), and image quality measures (PSNR, SSIM) to validate effectiveness.

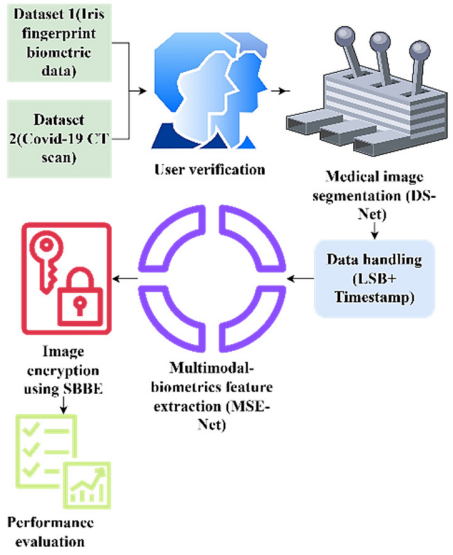


Fig. 1. Overall architecture of the proposed model.

A. Data Collection

Multimodal Iris Fingerprint Biometric data was collected from [21]. This dataset is open-access, publicly available under Kaggle's standard dataset license, and anonymized to protect personal identity, complying with general ethical guidelines for biometric data research.

COVID-19 CT scan lesion segmentation data was also collected from [22]. All data are de-identified and open-source, collected under previously approved ethical protocols, allowing use for academic and research purposes without infringing patient privacy.

B. User Verification (OTP Authentication)

OTP generation can be mathematically expressed using:

$$OTP_{user} = f(K_{\text{secret}}, T_{\text{stamp}}) \quad (1)$$

where $f(\cdot)$ defines the cryptographic hash function, K_{secret} defines the pre-shared key between the user and system, and T_{stamp} refers to the current timestamp, ensuring temporal uniqueness. The following equation describes the OTP verification.

$$\text{Access} = \begin{cases} 1 & \text{if } OTP_{input} = OTP_{user} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

OTP authentication is the initial security level.

C. Medical Image Segmentation (DS-Net)

DS-Net is a deep learning U-shaped network aimed at segmenting medical images into accurate lesion and non-lesion regions. Initially, the CT image is preprocessed for normalization and resizing to ensure consistent input for DS-Net. The input data are normalized using:

$$I_{\text{norm}} = \frac{I_{\text{medical}} - \mu}{\sigma} \quad (3)$$

where I_{medical} defines the original medical CT image, μ and σ represent the mean and standard deviation of pixel intensities and I_{norm} is the normalized image.

CalibNeXt is a convolutional encoder that improves feature representation by adaptively refining spatial and channel information for accurate lesion detection. It integrates a Spatial-Channel Calibration (SCC) module, which performs multi-scale feature extraction using adaptive weighting. SCC enhances important lesion regions and suppresses noise, improving segmentation accuracy in complex medical images. CalibNeXt uses SCC for multi-scale feature extraction, enhancing lesions:

$$F_{\text{local}} = \text{CalibNeXt}_{\text{SCC}}(I_{\text{norm}}) \quad (4)$$

where F_{local} defines the feature map representing local structures of the lesion. The Residual Swin Transformer (RST) captures global contextual information, such as the overall structure and shape of lesions, mathematically expressed using:

$$F_{\text{global}} = \text{RST}(I_{\text{norm}}) \quad (5)$$

where F_{global} is the feature map representing the global lesion context. Then, the extracted local and global features are fused through a Triplet Attention Mechanism (TAM), which emphasizes the most relevant lesion regions. Thus, the feature fusion process can be mathematically expressed as:

$$F_{\text{fused}} = \text{TAM}(F_{\text{local}}, F_{\text{global}}) \quad (6)$$

where the F_{fused} parameter encloses the attention-weighted feature map. The lesion mask prediction can be mathematically expressed as:

$$S_{\text{lesion}} = \text{Decoder}(F_{\text{fused}}) \quad (7)$$

where S_{lesion} defines the segmented lesion mask. The Segmentation loss function is given by:

$$\mathcal{L}_{\text{seg}} = \lambda_1 \mathcal{L}_{\text{Dice}}(S_{\text{lesion}}, S_{\text{gt}}) + \lambda_2 \mathcal{L}_{\text{CE}}(S_{\text{lesion}}, S_{\text{gt}}) \quad (8)$$

where S_{gt} represents the ground truth lesion mask, and λ_1, λ_2 are the weighting factors for each loss.

Figure 2 presents the DS-Net architecture, which employs a dual-encoder strategy for precise medical image segmentation. The first encoder, based on CalibNeXt, extracts local features using batch normalization, ReLU activation, and the SCC module, which enhances spatial and channel-wise representations through parallel branches and residual connections. The second encoder utilizes an RST to capture global contextual information via patch embedding, shifted window multi-head self-attention, and MLP layers. The local (F_{local}) and global (F_{global}) features are fused through the

TAM, enabling robust multi-scale feature integration and accurate lesion boundary detection.

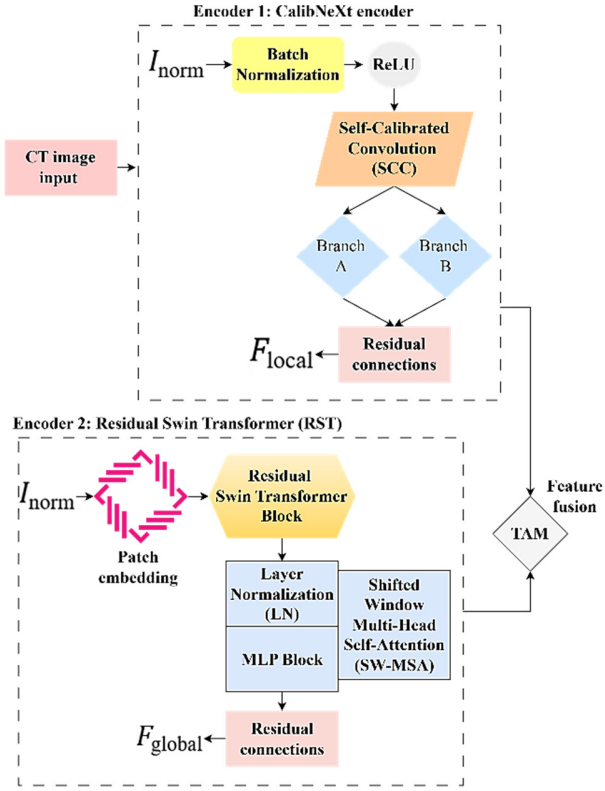


Fig. 2. Architecture of DS-Net.

D. Data Hiding (LSB+Timestamp)

From the previous step, the segmentation network produces two key outputs. The non-lesion region $I_{\text{non-lesion}}$ is obtained using:

$$I_{\text{non-lesion}} = I_{\text{medical}} - S_{\text{lesion}} \quad (9)$$

where I_{medical} is the original medical CT image.

The embedding process is represented mathematically by:

$$I_{\text{marked}}(x,y) = I_{\text{non-lesion}}(x,y) \oplus \left(S_{\text{lesion}}(x,y) \parallel T_{\text{stamp}}\right) \quad (10)$$

where $I_{\text{marked}}(x,y)$ defines the pixel intensity of the marked (watermarked) image, $\oplus$ represents the bitwise embedding operation (LSB substitution), $\parallel$ is the concatenation operator, $S_{\text{lesion}}(x,y)$ is the lesion bit pattern at pixel (x, y) and T_{stamp} represents the timestamp value appended for temporal uniqueness. The timestamp is taken from the system clock, and a hash is created using a cryptographic function to ensure that the data cannot be changed.

$$T_{\text{stamp}} = h(\text{DateTime}_{\text{current}} \parallel K_{\text{secret}}) \quad (11)$$

where $h(\cdot)$ is a secure hash function (for instance, SHA-256), and K_{secret} is the secret key shared during OTP verification.

Finally, after embedding all the lesion and timestamp bits, the complete marked image is reconstructed using:

$$I_{\text{marked}} = \bigcup_{x,y} I_{\text{marked}}(x,y) \quad (12)$$

where $(\bigcup_{x,y})$ denotes the union (pixel-wise aggregation) of all embedded blocks.

BioMedShield securely hides lesion data and timestamps in non-lesion areas using LSB, ensuring traceable, lossless, and visually intact medical images.

E. Multimodal Biometric Feature Extraction (MSE-Net)

Before feature extraction, biometric data undergo normalization to maintain a consistent intensity distribution:

$$I_{\text{bio-norm}} = \frac{I_{\text{biometric}} - \mu_{\text{bio}}}{\sigma_{\text{bio}}} \quad (13)$$

where $I_{\text{biometric}}$ defines the raw biometric image (iris or face), μ_{bio} and σ_{bio} are the mean and standard deviation of pixel intensities, and $I_{\text{bio-norm}}$ defines the normalized biometric image. Feature encoding can be mathematically expressed as:

$$K_{\text{enc}} = \text{Encode}(F_{\text{bio}}, K_{\text{secret}}, D_{\text{addr}}) \quad (14)$$

where K_{enc} defines the encoded encryption key vector, F_{bio} is the biometric feature vector, K_{secret} represents the sender's secret authentication key and D_{addr} states the unique device identifier. To keep the encoding function non-linear and unique, one can consider that the function represents a composition of cryptographic operations on parameters that were given as inputs, mathematically expressed as:

$$K_{\text{enc}} = h((F_{\text{bio}} \oplus K_{\text{secret}}) \parallel D_{\text{addr}}) \quad (15)$$

where $h(\cdot)$ is a secure cryptographic hash function, $\oplus$ is the bitwise XOR operation combining the biometric features and secret key, and $\parallel$ defines the concatenation operator, combining the result with the device address.

Figure 3 illustrates the MSE-Net architecture designed for multimodal biometric feature extraction and secure key generation. Biometric inputs and marked medical images are first normalized and passed through a multiscale input conversion block to capture features at different resolutions. These features are processed using an EfficientNet backbone, which includes convolutional layers, inverted residual blocks, and squeeze-and-excitation modules for efficient representation learning. A dilated convolution block with residual and concatenation operations further refines feature extraction. The resulting biometric and annotated image features are fused to form a comprehensive feature vector ($F_{\text{bio-total}}$), which is then encoded to generate a secure encryption key (K_{enc}).

F. Image Encryption Using SBBE

SBBE secures lesion-marked medical images using user- and device-specific biometric keys, refined via evolutionary optimization, ensuring high-quality, tamper-resistant encryption with confidentiality, integrity, and non-repudiation.

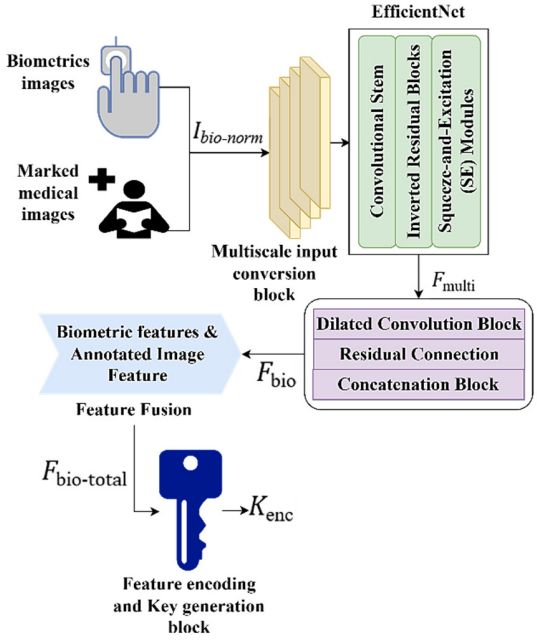


Fig. 3. Architecture of MSE-Net.

1) Key Initialization and Population Setup

Initially, candidate keys are set up before the encryption takes place. The initial population of candidate keys can be mathematically expressed as:

$$\mathcal{P}^{(0)} = \{K_i^{(0)} \mid K_i^{(0)} = f_{\text{init}}(K_{\text{enc}}, rand_i),\ i = 1,2,\dots,N\} \quad (16)$$

where $\mathcal{P}^{(0)}$ is the initial population of candidate keys, $K_i^{(0)}$ defines the i-th candidate key in generation 0, $f_{\text{init}}(\cdot)$ is the key initialization function (e.g., XOR, hash-based mixing), K_{enc} is the encoded biometric key, $rand_i$ defines the random seed for the candidate i, ensuring population diversity, and the population size is denoted as N (20–50).

2) Fitness Evaluation

Encryption maximizes randomness, minimizes correlation, and enhances sensitivity using the candidate key given as:

$$C_i = \text{Enc}_{\text{SBBE}}(I_{\text{sample}}, K_i^{(g)}) \quad (17)$$

Thus, the fitness computation can be expressed as:

$$\text{Fitness}\left(K_i^{(g)}\right) = f\left(H(C_i), r(C_i), \text{NPCR}(C_i), \text{UACI}(C_i)\right) \quad (18)$$

where C_i is the ciphertext obtained using the candidate key $K_i^{(g)}$, I_{sample} is the representative image block or down-sampled image, $H(C_i)$ is the Shannon entropy of ciphertext (higher indicates more randomness), $r(C_i)$ is the correlation coefficient between adjacent pixels (lower is better), NPCR stands for Number of Pixels Change Rate, which indicates plaintext sensitivity, and UACI stands for Unified Average Changing Intensity, which is a measure of intensity variation.

3) Global Exploration

Global exploration is an essential part of the SBBE system to create diversity in the candidate key population.

- Random recombination: This operator combines two candidate keys K_a and K_b from the current population to produce a new offspring key, and it is mathematically generated using:

$$K_{\text{new}} = K_a^{(g)} \oplus \text{Rot}(K_b^{(g)}, r), K_a, K_b \in P^{(g)} \quad (19)$$

where $\text{Rot}(\cdot)$ denotes bitwise rotation by r positions.

- Large mutation: Randomly selected blocks of bits within a key are flipped with a specified probability p_{mut}:

$$K_{\text{mut}}[B] = K[B] \oplus M, M \sim \{0,1\}^{|B|} \quad (20)$$

Large mutation introduces abrupt variations, allowing the system to explore regions of the key space that may not be reachable through recombination alone.

- Random injection: Portions of a candidate key are replaced with hashed seeds derived from the encoded biometric key and a random number:

$$K_{\text{inj}} = \text{Replace}(K, pos, \text{Trunc}(\text{h}(K_{\text{enc}} \parallel rand))) \quad (21)$$

$K_a^{(g)}$ and $K_b^{(g)}$ are the parent keys and the function $\text{Rot}(\cdot, r)$ refers to performing the bitwise rotation of a key by r positions. Random injection replaces portions of a key at a specified position pos with a truncated cryptographic hash $\text{h}(\cdot)$ of the encoded biometric key concatenated with a random number $rand$, ensuring both randomness and biometric/device dependency.

4) Exploitation / Local Search

Local search methods are used to further enhance the best candidate keys derived from the top-performing ones.

5) Selection and Replacement

Top N keys are chosen after the fitness evaluation to constitute the next generation. This can be mathematically expressed using:

$$P^{(g+1)} = \text{Top-}N \text{ keys based on Fitness} \quad (22)$$

$P^{(g+1)}$ defines the population for the next generation, and the fitness can be computed as a metric that combines entropy, correlation, NPCR, and UACI.

6) Termination

The best key K^* is chosen for the final encryption, generated using:

$$K^* = \arg\max \text{Fitness}(K_i) \quad (23)$$

Here, K^* is the optimized key selected for SBBE encryption.

7) Image Encryption

The final marked medical image is encrypted using the optimized key K^*. The encrypted images are expressed as:

$$C_{\text{final}} = \text{SBBE}_{\text{Encrypt}}(I_{\text{marked}}, K^*) \quad (24)$$

where C_{final} defines the ciphertext of the marked medical image, I_{marked} is the marked image after data hiding, and K^* represents the best encryption key obtained from evolutionary optimization.

SBBE is a secure medical image encryption method that uses biometric-based keys and evolutionary optimization to generate unique session keys. By combining user-specific biometrics with optimized Blowfish encryption, it ensures strong protection against brute-force, statistical, and differential attacks in IIoT healthcare systems. Figure 4 details the sequential steps involved in the image encryption process through the SBBE (Swarm-Based Block Encryption) method.

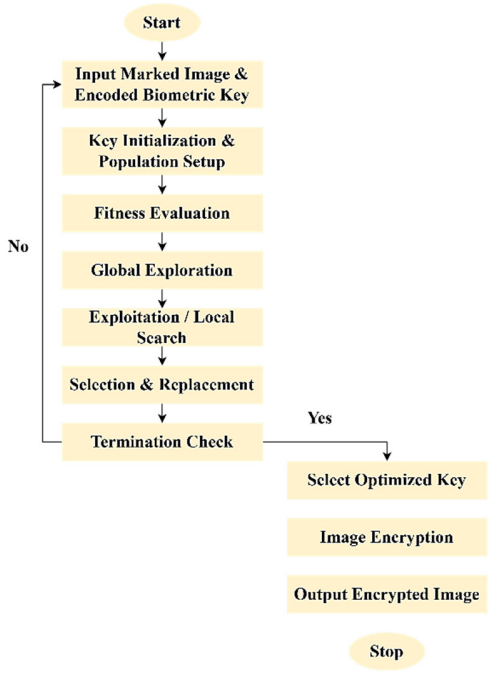


Fig. 4. Flowchart for SBBE-based image encryption.

TABLE II. EXPERIMENTAL SETUP FOR BIOMEDSHIELD FRAMEWORK EVALUATION

Parameter	Configuration / Value
Segmentation network	DS-Net
Biometric feature extractor	MSE-Net
Encryption algorithm	SBBE
DS-Net optimizer	Adam
DS-Net learning rate	0.001
DS-Net batch size	16
DS-Net epochs	20
MSE-Net optimizer	Adam
MSE-Net learning rate	0.0005
MSE-Net batch size	16
MSE-Net epochs	25
Evaluation metrics	Accuracy, Precision, Recall, F1-Score, Specificity, Dice Coefficient
Hardware	GPU-enabled workstation (NVIDIA RTX 3090)

III. RESULTS AND ANALYSIS

A. Experimental Setup

The proposed BioMedShield framework was experimentally evaluated using a comprehensive multimodal dataset and the setup shown in Table I.

B. Metrics Analysis

Table III summarizes the key performance metrics used to evaluate the proposed BioMedShield framework.

TABLE III. PERFORMANCE METRICS

Metric	Description	Formula
Accuracy (%)	Overall correct predictions	$Accuracy = \frac{tp + tn}{tp + fp + tn + fn} \times 100$
Precision (%)	Correct positive predictions	$Precision = \frac{tp}{tp + fp} \times 100$
Recall (%)	True positive rate	$Recall = \frac{tp}{tp + fn} \times 100$
F1-Score (%)	Balance precision and recall	$F1 - score = 2 \times \frac{precision \times Recall}{precision + Recall} \times 100$
Specificity (%)	True negative rate	$Specificity = \frac{tn}{tn + fp} \times 100$
Dice Coefficient (%)	Overlap similarity measure	$Dice = \frac{2 \times \lvert Prediction \cap Ground\ truth \rvert}{\lvert Prediction \rvert + \lvert Ground\ truth \rvert} \times 100$

C. Performance Analysis

The baseline models—TransUNet [23], U-Net [24], Attention U-Net [24], and SegNet [3]—were selected because they represent widely used and state-of-the-art medical image segmentation architectures. U-Net and SegNet serve as classical encoder-decoder benchmarks, while Attention U-Net enhances lesion localization through attention mechanisms. TransUNet was included as a modern CNN-Transformer hybrid for comparison with recent transformer-based methods. All models were implemented and trained on the same CT dataset using identical preprocessing, training-testing splits, input resolution, optimization settings, and evaluation metrics. The choice of these baselines is justified as follows:

- U-Net and SegNet: Strong classical benchmarks for encoder–decoder medical segmentation.
- Attention U-Net: Represents attention-based enhancement for improved lesion focus.
- TransUNet: Represents recent transformer-based global context learning.

Thus, the comparison covers classical CNN, attention-enhanced CNN, and transformer-based architectures, demonstrating the effectiveness of DS-Net across different methodological categories. Table IV shows that DS-Net outperforms TransUNet, Attention U-Net, Secure Encryption, and SegNet, achieving 98.7% recall for robust, secure medical image segmentation [26]. Figure 5 shows that MSE-Net is better than all the baseline models in biometrics classification. Figure 6 shows that DS-Net was superior in lesion segmentation.

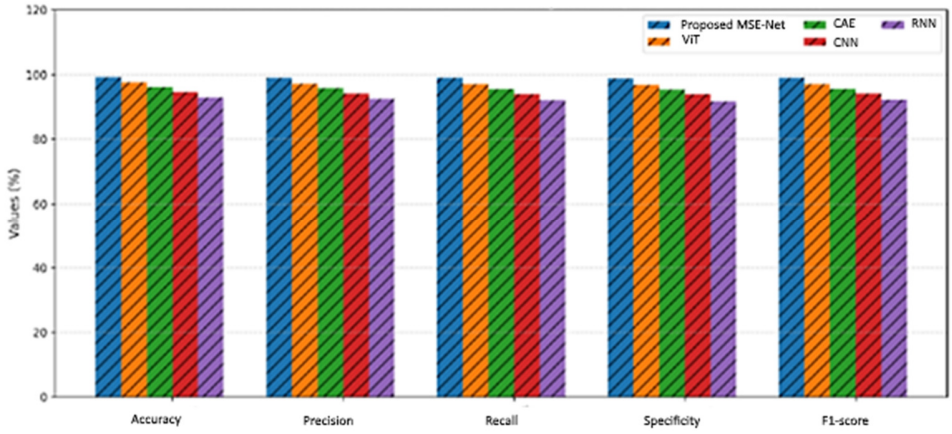


Fig. 5. Comparative performance analysis of the proposed MSE-Net and baseline models for biometric classification.

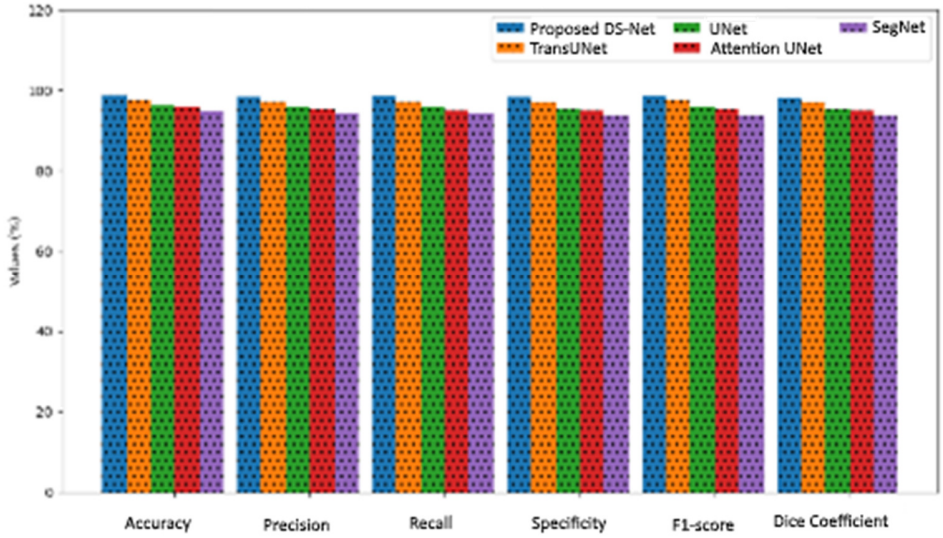


Fig. 6. Lesion segmentation performance comparison between DS-Net and baseline models.

TABLE IV. PERFORMANCE COMPARISON OF THE PROPOSED WITH EXISTING MODELS

	Proposed DS-Net	**TransUNet [23]**	**U-Net [24]**	**Attention U-Net [25]**	**SegNet [3]**
Accuracy	98.8	97.6	96.4	95.8	94.9
Precision	98.5	97.3	96	95.4	94.3
Recall	98.7	97.2	95.8	95.1	94.1
Specificity	98.4	97.1	95.5	95	93.8
F1-score	98.6	97.4	95.9	95.3	94
Dice Coefficient	98.3	97	95.4	95	93.7

Figure 7 shows the confusion matrix for the MSE-Net, which demonstrates that the network was able to classify biometrics accurately to a great extent for the ten classes. Figure 8 illustrates a comparison between unimodal and multimodal biometric recognition systems. Figure 9 shows an analysis of encryption and decryption time (ms) with respect to increasing image size (KB), clearly highlighting the scalability and superior efficiency of the proposed SBBE algorithm.

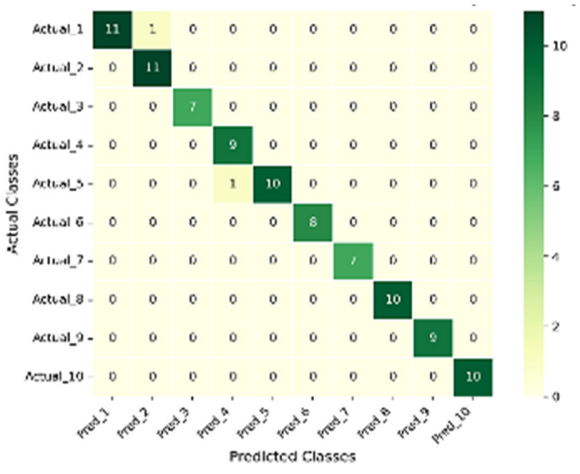


Fig. 7. Confusion matrix on biometric classification (MSE-Net).

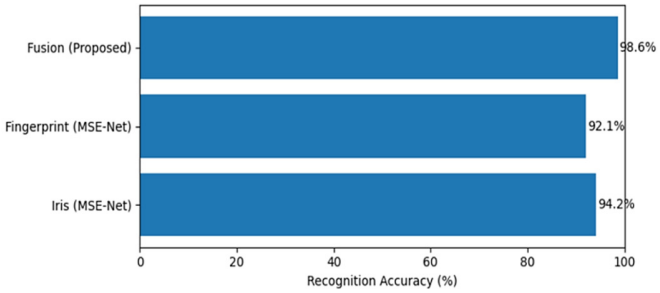


Fig. 8. Biometric recognition: Unimodal vs. multimodal fusion.

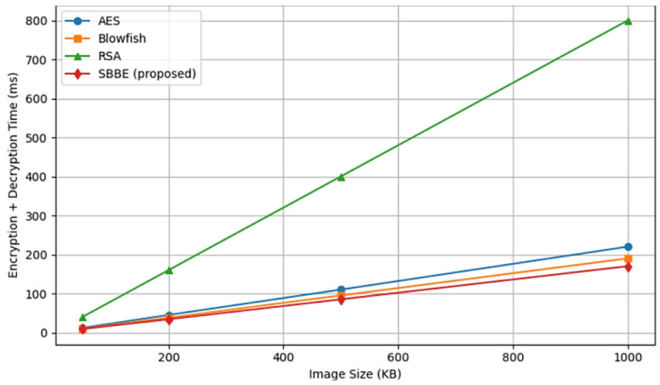
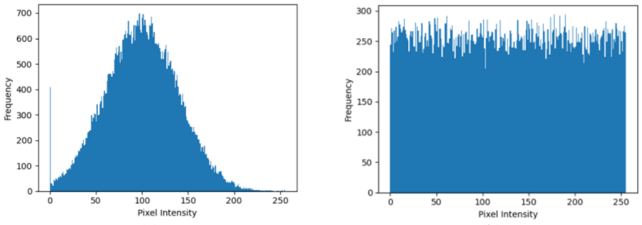


Fig. 9. Encryption performance vs image size.

Figure 10(a) presents the original medical image histogram, where the pixel intensities are unevenly distributed, with most of them concentrated within the medium intensity range. In contrast, Figure 10(b) depicts the histogram of the encrypted image, which exhibits an almost perfectly uniform distribution across the full range of pixel intensities (0 to 255).



(a) (b)

Fig. 10. (a) Original image histogram, (b) Encrypted image histogram.

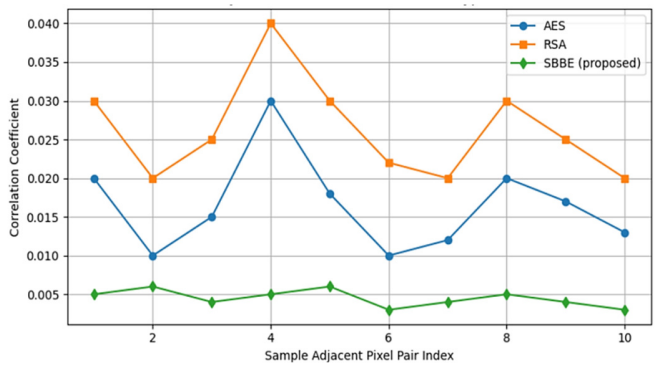


Fig. 11. PSNR and SSIM for decrypted images.

Figure 11 illustrates the correlation coefficients of adjacent pixels for images encrypted using various algorithms. Figure 12 illustrates the Receiver Operating Characteristic (ROC) curves. Figure 13 shows that the proposed SBBE algorithm achieves the fastest encryption performance, completing the task in just 0.67 s.

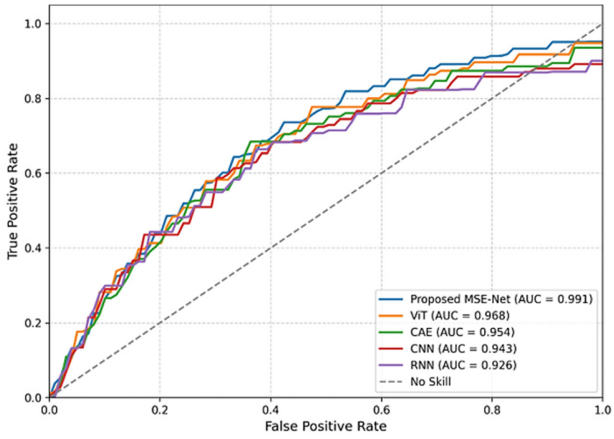


Fig. 12. ROC curves for biometric authentication.

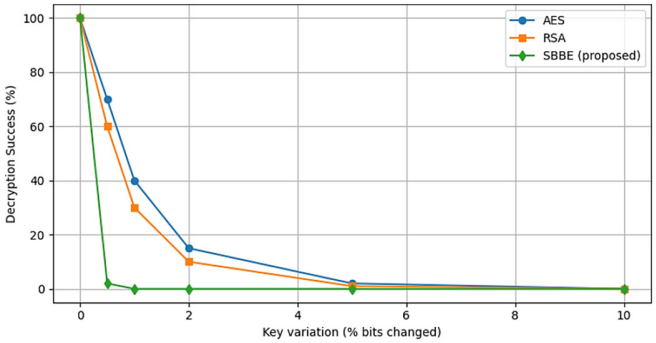


Fig. 13. Key sensitivity / Differential attack resistance.

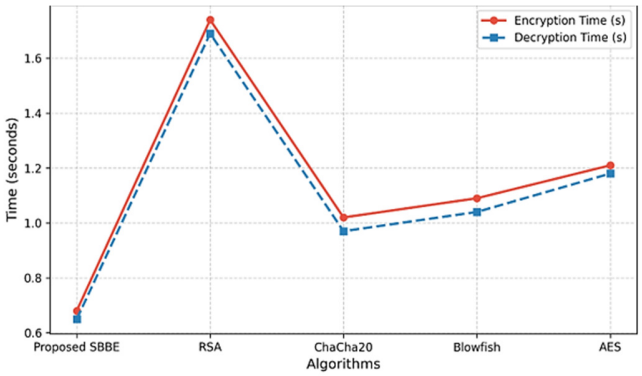


Fig. 14. Encryption vs decryption time comparison.

Figure 15 illustrates the training and validation accuracy of the model over 20 epochs. Figure 16 shows the training and validation loss over 20 epochs. Figure 17 compares the security levels (in bits) of five different cryptographic algorithms.

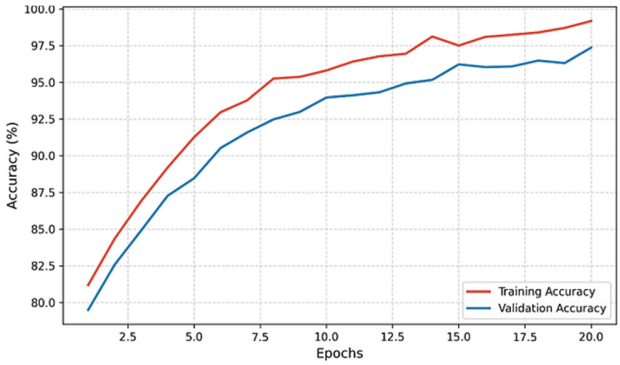


Fig. 15. Training and validation accuracy

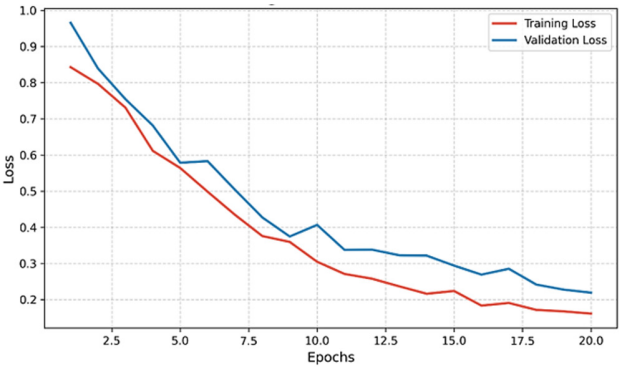


Fig. 16. Training and validation loss.

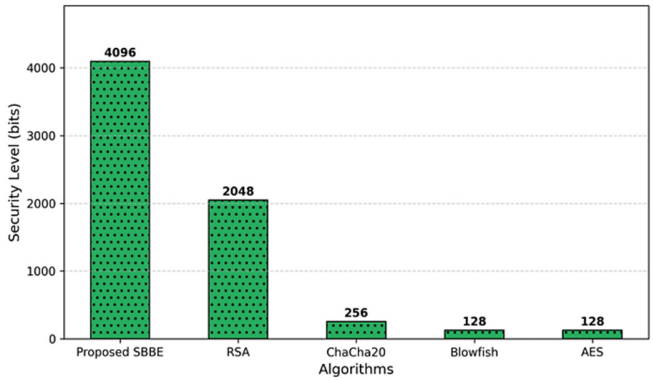


Fig. 17. Security level comparison.

Table V features a detailed comparison of the proposed DS-Net model with several state-of-the-art models in terms of various performance measures such as Accuracy, Precision, Recall, Specificity, F1-Score, and Dice Coefficient. As also shown in Figure 18, the Proposed DS-Net far outperforms all existing state-of-the-art models in these metrics.

D. Discussion

BioMedShield ensures secure medical image sharing using multimodal biometrics, dual-network deep learning, and SBBE encryption, but GPU dependence, latency, and scalability remain challenges. SBBE enhances medical image security with biometric, evolutionary, and Blowfish encryption, but may be computationally intensive and biometric-dependent.

TABLE V. COMPARISON OF THE PROPOSED MODEL WITH THE STATE-OF-THE-ART MODELS

Model	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-score (%)	Dice Coefficient (%)
DL [13]	94.2	93.8	93.5	93.1	93.6	93.2
DeepENC [14]	95.4	95.1	94.8	94.5	94.9	94.6
PPAF-BS [15]	96.1	95.7	95.9	95.3	95.8	95.4
Coupled Chaotic Mapping [16]	96.8	96.4	96.5	96.1	96.3	96
LSB [17]	97.1	96.8	96.7	96.3	96.6	96.2
CNN + RNN Hybrid [18]	97.5	97.3	97	96.8	97.1	96.7
DNN-ChOA [19]	97.8	97.6	97.4	97.1	97.5	97.2
SEFF-CNN-BO [20]	98.1	98	97.8	97.5	97.9	97.6
Proposed DS-Net	98.8	98.5	98.7	98.4	98.6	98.3

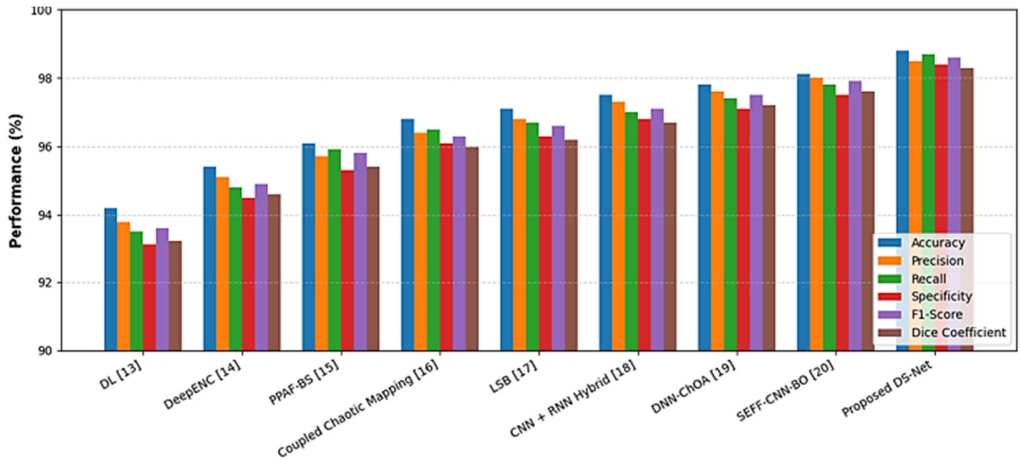


Fig. 18. Graphical representation of the comparison between the state of the art and the proposed model.

IV. CONCLUSION

This study addressed a critical gap in IIoT-enabled healthcare systems, where existing solutions focus on encryption or biometric authentication independently but lack an integrated, end-to-end framework that simultaneously ensures secure access control, data confidentiality, and structural preservation of medical images. To bridge this limitation, the proposed BioMedShield framework introduced a unified deep learning-based multimodal biometric protection system that combines medical image segmentation, biometric feature extraction, data hiding, and encryption within a single architecture. The novelty of this work lies in three key aspects. First, the proposed DS-Net integrates CalibNeXt-based local feature learning with an RST for global contextual understanding, enabling precise lesion segmentation while preserving clinically relevant structures. Second, sensitive lesion information is securely embedded into non-lesion regions using an adaptive LSB-based strategy, ensuring data integrity without affecting diagnostic quality. Third, the SBBE mechanism generates strong encryption keys using multimodal biometric features and device-specific information, providing identity-bound protection suitable for IIoT environments. The experimental results demonstrated the effectiveness of the proposed approach, where DS-Net achieved 98.8% accuracy, 98.7% recall, and a 98.3% Dice coefficient, outperforming standard models such as TransUNet, U-Net, Attention U-Net, and SegNet. The integrated security pipeline ensured robust protection against unauthorized access while maintaining high image quality and segmentation reliability, making the system practical for real-world telemedicine and remote diagnostic applications. Overall, BioMedShield contributes a scalable, secure, and intelligent medical image protection framework that enhances privacy, authentication reliability, and data integrity in distributed healthcare environments. By combining deep learning-based segmentation with multimodal biometric encryption and secure data hiding, the proposed system provides a comprehensive solution for safeguarding sensitive medical data in next-generation IIoT healthcare systems.

BioMedShield can evolve for IIoT healthcare by improving edge-device efficiency for real-time processing, integrating additional biometrics such as iris or voice, and leveraging federated learning and blockchain for privacy, decentralization, and traceability. Supporting low-resolution, noisy, or heterogeneous images enhances robustness, expanding its role in secure, intelligent healthcare systems.

REFERENCES

[1] R. Mekala, "Deep Learning in the Cloud for Anomaly Detection Improves Medical Monitoring," *Journal of International Exercise Sciences*, vol. 2, no. 1, pp. 44–59, 2023.

[2] F. Castro, D. Impedovo, and G. Pirlo, "A Medical Image Encryption Scheme for Secure Fingerprint-Based Authenticated Transmission," *Applied Sciences*, vol. 13, no. 10, May 2023, https://doi.org/10.3390/app13106099.

[3] M. Singh, N. Baranwal, K. N. Singh, A. K. Singh, and H. Zhou, "Deep learning-based biometric image feature extraction for securing medical images through data hiding and joint encryption–compression," *Journal of Information Security and Applications*, vol. 79, Dec. 2023, Art. no. 103628, https://doi.org/10.1016/j.jisa.2023.103628.

[4] Z. Liu and R. Xue, "Medical Image Encryption using Biometric Image Texture Fusion," *Journal of Medical Systems*, vol. 47, no. 1, Nov. 2023, Art. no. 112, https://doi.org/10.1007/s10916-023-02003-5.

[5] W. El-Shafai, I. Almomani, A. Ara, and A. Alkhayer, "An optical-based encryption and authentication algorithm for color and grayscale medical images," *Multimedia Tools and Applications*, vol. 82, no. 15, pp. 23735–23770, June 2023, https://doi.org/10.1007/s11042-022-14093-3.

[6] N. Santos, B. Ghita, and G. L. Masala, "Medical Systems Data Security and Biometric Authentication in Public Cloud Servers," *IEEE Transactions on Emerging Topics in Computing*, vol. 12, no. 2, pp. 572–582, Apr. 2024, https://doi.org/10.1109/TETC.2023.3271957.

[7] J. Selvaraj, W. C. Lai, B. P. Kavin, K. C, and G. H. Seng, "Cryptographic Encryption and Optimization for Internet of Things Based Medical Image Security," *Electronics*, vol. 12, no. 7, Mar. 2023, https://doi.org/10.3390/electronics12071636.

[8] D. Zhang, L. Ren, M. Shafiq, and Z. Gu, "A Privacy Protection Framework for Medical Image Security without Key Dependency Based on Visual Cryptography and Trusted Computing," *Computational Intelligence and Neuroscience*, vol. 2023, no. 1, 2023, Art. no. 6758406, https://doi.org/10.1155/2023/6758406.

[9] A. S. Jamil, R. A. Azeez, and N. F. Hassan, "An Image Feature Extraction to Generate a Key for Encryption in Cyber Security Medical Environments," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 19, no. 01, pp. 93–106, Jan. 2023, https://doi.org/10.3991/ijoe.v19i01.36901.

[10] A. A. Alhussan, H. A. Abdallah, S. Alsodairi, and A. A. Ateya, "Hybrid Watermarking and Encryption Techniques for Securing Medical Images," *Computer Systems Science and Engineering*, vol. 46, no. 1, pp. 403–416, 2023, https://doi.org/10.32604/csse.2023.035048.

[11] A. Odeh and A. A. Taleb, "A Multi-Faceted Encryption Strategy for Securing Patient Information in Medical Imaging," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 14, no. 4, pp. 164–176, Dec. 2023, https://doi.org/10.58346/JOWUA.2023.I4.012.

[12] Y. Qin and B. Zhang, "Privacy-Preserving Biometrics Image Encryption and Digital Signature Technique Using Arnold and ElGamal," *Applied Sciences*, vol. 13, no. 14, July 2023, https://doi.org/10.3390/app13148117.

[13] K. N. Singh, N. Baranwal, A. K. Singh, A. K. Agrawal, and H. Zhou, "Using Multimodal Biometrics, Data Hiding, and Encryption for Secure Healthcare Imaging System," *IEEE Transactions on Consumer Electronics*, vol. 71, no. 2, pp. 6711–6718, Feb. 2025, https://doi.org/10.1109/TCE.2024.3438356.

[14] K. N. Singh, N. Baranwal, O. P. Singh, and A. K. Singh, "DeepENC: Deep Learning-Based ROI Selection for Encryption of Medical Images Through Key Generation With Multimodal Information Fusion," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 3, pp. 6149–6156, Dec. 2024, https://doi.org/10.1109/TCE.2024.3406963.

[15] S. B. Jadhav, N. K. Deshmukh, and S. B. Pawar, "Robust Authentication System with Privacy Preservation for Hybrid Deep Learning-Based Person Identification System Using Multi-Modal Palmprint, Ear, and Face Biometric Features," *International Journal of Image and Graphics*, vol. 25, no. 05, Sept. 2025, Art. no. 2550049, https://doi.org/10.1142/S0219467825500494.

[16] H. Xie, Y. Zhang, J. Bian, and H. Zhang, "A secure and privacy-preserving technique based on coupled chaotic system and plaintext encryption for multimodal medical images," *Multimedia Tools and Applications*, vol. 84, no. 20, pp. 22701–22726, June 2025, https://doi.org/10.1007/s11042-024-19956-5.

[17] G. Latif, J. Alghazo, N. Mohammad, S. E. Abdelhamid, G. B. Brahim, and K. Amjad, "A Novel Fragmented Approach for Securing Medical Health Records in Multimodal Medical Images," *Applied Sciences*, vol. 14, no. 14, July 2024, https://doi.org/10.3390/app14146293.

[18] S. Vatchala *et al.*, "Multi-Modal Biometric Authentication: Leveraging Shared Layer Architectures for Enhanced Security," *IEEE Access*, vol. 13, pp. 28029–28041, 2025, https://doi.org/10.1109/ACCESS.2025.3534223.

[19] F. Sayeed, K. R. Ahmed, and S. M. Swamy, "Development of a multimodal biometric recognition system with feature optimization and

deep learning," *Multimedia Tools and Applications*, vol. 84, no. 31, pp. 38399–38422, Sept. 2025, https://doi.org/10.1007/s11042-025-20709-1.

[20] M. S. Nair and S. Dharan, "Design a Multimodal Biometric Based Protection System by Generation of a Revocable Cryptographic Key Using Separately Extracted Feature Fusion-based Convolutional Neural Network With Bat Optimization." Research Square, Aug. 29, 2024, https://doi.org/10.21203/rs.3.rs-4853162/v1.

[21] "Multimodal_Iris_Fingerprint_Biometric_data." Kaggle, [Online]. Available: https://www.kaggle.com/datasets/ninadmehendale/multimodal-iris-fingerprint-biometric-data.

[22] "COVID-19 CT scan lesion segmentation dataset." Kaggle, [Online]. Available: https://www.kaggle.com/datasets/maedemaftouni/covid19-ct-scan-lesion-segmentation-dataset.

[23] S. Bian, X. Xu, W. Jiang, Y. Shi, and T. Sato, "BUNET: Blind Medical Image Segmentation Based on Secure UNET," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 612–622, https://doi.org/10.1007/978-3-030-59713-9_59.

[24] S. Subathra and V. Thanikaiselvan, "Enhanced security for medical images using a new 5D hyper chaotic map and deep learning based segmentation," *Scientific Reports*, vol. 15, no. 1, July 2025, Art. no. 22628, https://doi.org/10.1038/s41598-025-04906-4.

[25] H. Farah, A. Bennourm, H. Soltani, M. Nahas, R. R. Marie, and M. Al-Sarem, "Attention U-Net for Precision Skeletal Segmentation in Chest X-Ray Imaging: Advancing Person Identification Techniques in Forensic Science," *Computers, Materials and Continua*, vol. 85, no. 2, pp. 3335–3348, Sept. 2025, https://doi.org/10.32604/cmc.2025.067226.

[26] Q. Chen, H. Li, S. B. Ariffin, and N. A. B. Mustapa, "A Comprehensive Study on the Homomorphic Encryption for Secure Image Data Processing," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 21783–21790, Apr. 2025, https://doi.org/10.48084/etasr.10007.