IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Enhanced Document Security Throught Biometric Watermarking And Machine Learning

[1]Naren Rakshith KV, [2]Vishva Kiran RC, [3]Ravitej Arjun Kakhandaki, [4]Rakshita G Sataraddi, [5]Samrat Singh
[1]Student, [2]Assistant Professor, [3]Student, [4]Student, [5]Student,
[1]Department of CSE, [2]Department of CSE, [3]Department of CSE, [4]Department of CSE, [5]Department of CSE,
[1]KSIT, Bengaluru, India, [2]KSIT, Bengaluru, India, [3]KSIT, Bengaluru, India, [4]KSIT, Bengaluru, India,
[5]KSIT, Bengaluru, India

Abstract: As virtual statistics turns into an increasing number of popular ensuring strong safety for sensitive files is critical this studies introduces a complicated protection framework that mixes biometric watermarking with device mastering to establish a tamper-resistant and adaptive protection system by way of encoding intricate iris and fingerprint patterns the usage of a custom designed rubiks cube encryption algorithm the method creates a comfy embedded watermark that is tremendously proof against manipulation in parallel convolutional neural networks CNNs examine and authenticate biometric statistics permitting real-time detection of spoofing tries and unauthorized changes the adaptive gaining knowledge of functionality of CNNs lets in the system to refine its detection accuracy through the years strengthening its resilience against rising threats this precise integration of encryption and shrewd pattern recognition gives extensive improvements in file security with ability packages in sectors which include healthcare finance and authorities wherein records integrity and authentication are paramount.

Keywords - Biometric Watermarking, Document Security, Rubik Encryption, Convolutional Neural Networks (CNN), Machine Learning, Iris and Fingerprint Fusion, Zero-bit Watermarking, Authentication, Spoofing Detection, Fraud Detection.

I. Introduction

The digital age has brought forth a critical need for fortified document security, as conventional methods struggle to with- stand the evolving landscape of cyber threats. This research introduces a novel security paradigm, intertwining biometric watermarking with adaptive machine learning, to forge an unbreakable link between a document and its rightful owner. By harnessing the unique, immutable characteristics of human biometrics—specifically, iris and fingerprint patterns—this system generates an embedded, concealed watermark that acts as a tamper-evident seal.

The generation of this watermark involves a multi-layered encryption process, deviating from standard cryptographic approaches. A customized permutation algorithm, inspired by the principles of the Rubik’s cube, is employed to scramble and embed biometric data, creating a highly resilient watermark that is virtually impossible to replicate or manipulate. To ensure real-time authentication and detect any unauthorized modifications, the system leverages Convolutional Neural Net- works (CNNs). These neural networks are trained to discern subtle variations between genuine and fabricated biometric data, providing a dynamic defense against spoofing and fraudulent access attempts.

What sets this system apart is its capacity for continuous refinement. The CNNs, through iterative learning, adapt to new data patterns, enhancing their ability to identify emerging threats. This adaptive intelligence ensures that the security framework remains robust and effective over time. By merging complex encryption with intelligent pattern recognition, this research aims to establish a new standard for document security, with applications spanning critical sectors such as healthcare, finance, and governance, where data integrity and authenticity are paramount. This unique combination of a permutation-based encryption and an adaptive CNN for real time spoofing detection offers a level of security that distinguishes it from other methods.

II. LITERATURE SURVEY

In recent years, the integration of biometric modalities, such as iris and fingerprint images, has gained considerable attention in the development of secure authentication systems. With the increasing need for tamper-proof biometric identification, researchers have explored various techniques for effective biometric feature extraction, fusion, and secure data embedding. Advanced encryption methods, particularly the Advanced Encryption Standard (AES), are widely used to protect embedded data. Additionally, Convolutional Neural Networks (CNNs) have become a powerful tool for learning discriminative features from biometric modalities, thereby improving the accuracy of authentication mechanisms. This section provides an overview of relevant studies in the fields of biometric watermarking, multimodal fusion, secure encryption, and deep learning-based authentication.

Shukla et al. [1] introduced a methodology that can utilize any image in JPG (JPEG) or PNG format, along with the text as input, to produce the text embedded within the image as output. This paper discusses retrieving the original text on the receiver's side without any data loss. Two classical encryption techniques, specifically Substitution and Transposition ciphers, are employed to enhance the security of transmitted text. Albkosh et al. [2] presented an efficient and secure digital watermarking system based on discrete wavelet transform (DWT) and AES encryption to ensure the integrity and confidentiality of medical images. Their method achieved high imperceptibility and resistance to image processing attacks. Verma et al. [3] explored a technique for multimodal biometric fusion using iris and fingerprint images by aligning them into a single composite biometric template. The fused biometric image was then encrypted using AES to establish a secure authentication scheme. Saini and Dutta [4] developed a secure biometric watermarking framework combining iris and fingerprint im- ages. The watermark was embedded using SVD-DWT and secured with 128-bit AES encryption to ensure confidentiality and robustness against common signal attacks. Yadav et al. [5] investigated the application of Convolutional Neural Networks (CNNs) to iris and fingerprint datasets to improve biometric identification accuracy. The fused CNN model outperformed individual models by learning better feature representations from the combined modalities. Raja et al. [6] proposed a watermarking-based biometric security system using discrete cosine transform (DCT) for embedding fingerprint features into iris images. The system ensured minimal distortion to the host image while maintaining strong resilience under geometric and compression attacks. Mohammed et al. [7] implemented an AES-based encryption system for multimodal biometric data to enhance data protection in cloud environments. Their fusion method maintained biometric integrity while improving computational efficiency. Gupta et al. [8] designed a hybrid model using DWT, SVD, and CNNs for biometric feature extraction and classification. The model demonstrated high precision in authenticating users through fused biometric data, especially under noisy conditions. Kumar and Rathi [9] proposed a lightweight CNN architecture for iris and fingerprint classification. Their system was trained on normalized datasets and proved effective in extracting robust features even from low-resolution biometric images. Ali and Farooq [10] introduced a system for secure watermark generation using entropy-based key generation for AES encryption. Their system showed improved security levels against brute- force and known-plaintext attacks while embedding biometric details in grayscale images.

III. PROPOSED SYSTEM

The proposed system introduces a unique, multi-layered biometric watermarking framework designed to enhance the security and authenticity of sensitive documents. This solution integrates cryptographic techniques and deep learning to create a tamper-resistant, intelligent authentication mechanism. The architecture comprises five core components: biometric encryption using the AES algorithm, secure biometric watermark generation, authentication via Convolutional Neural Network (CNN), a real-time notification system, and adaptive learning for continuous security enhancement. Additionally, the system employs one-time password (OTP)-based access control for watermark creation and removal, adding an extra layer of verification.

A. Biometric Encryption using AES Algorithm:

The encryption of iris and fingerprint biometric images using the Advanced Encryption Standard (AES) is funda- mental to this system. AES, a symmetric encryption algorithm, is widely recognized for its strength and efficiency. The biometric images undergo preprocessing before being encrypted using a 128-bit or 256-bit AES key, ensuring that the raw biometric content remains inaccessible even if data transmission is intercepted. The encryption process is defined by:

$$C = E_K(P) \qquad (1)$$

Where:

- C is the ciphertext (encrypted image),
- EK denotes the AES encryption function with key K
- P is the original biometric image (iris or fingerprint).

This encryption step ensures data confidentiality and establishes a secure foundation for watermark embedding

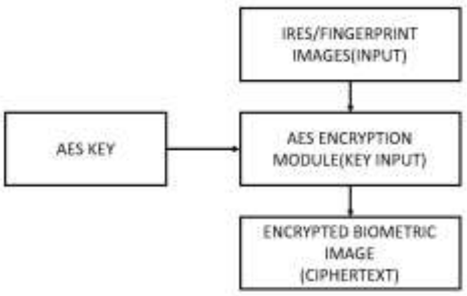
IRES/FINGERPRINT IMAGES(INPUT)
AES KEY
AES ENCRYPTION MODULE(KEY INPUT)
ENCRYPTED BIOMETRIC IMAGE (CIPHERTEXT)

Fig. 1: Biometric Encryption Flowchart

B. Biometric Watermark Generation:

Following encryption, the system generates a biometric watermark by fusing the encrypted iris and fingerprint images. This fusion process can utilize simple averaging or more sophisticated techniques like Discrete Wavelet Transform (DWT) or Principal Component Analysis (PCA) to maintain fidelity and robustness. The result is a composite watermark that retains identifiable features from both modalities while being resistant to reverse engineering. This watermark is subsequently embedded into digital documents using image watermarking techniques such as Least Significant Bit (LSB) substitution or frequency-domain embedding.

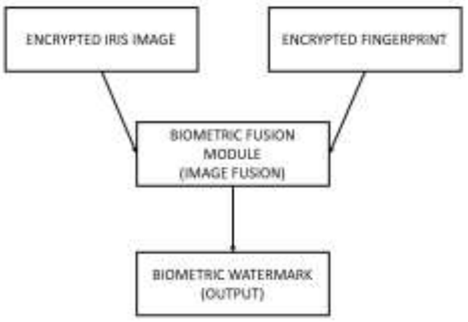
ENCRYPTED IRIS IMAGE
ENCRYPTED FINGERPRINT
BIOMETRIC FUSION MODULE (IMAGE FUSION)
BIOMETRIC WATERMARK (OUTPUT)

Fig. 2: Biometric Watermark Generation Flowchart

C. Convolutional Neural Network (CNN) for Authentication:

A critical component of the system is the use of a Convolutional Neural Network (CNN) for biometric authentication. The CNN is trained on a comprehensive dataset comprising both genuine and forged biometric watermarks. It extracts hierarchical features from the watermark and learns to classify whether a watermark is authentic. The CNN architecture includes convolutional layers, pooling layers, and fully connected layers, optimized for feature recognition and binary classification. The output of the network can be interpreted as:

$$Y = \sigma(Wx + b) \qquad (2)$$

Where:

- X represents the input watermark,
- W and b are weights and biases,
- σ is the activation function,
- Y is the classification output.

D. Adaptive Learning and Continuous Improvement:

To fortify the system against future threats, an adaptive learning mechanism is integrated into the CNN model. The model is retrained over time on new watermark patterns, including both legitimate and adversarial examples, enabling the CNN to refine its accuracy and maintain resilience against emerging spoofing techniques and adversarial attacks.

IV. IMPLEMENTATION AND METHODOLOGY

The presented biometric watermarking system is implemented through a sequential and modular approach, integrating cryptography, biometric image processing, deep learning, and secure access mechanisms to provide a robust and intelligent authentication platform.

A. Acquisition and Preprocessing of Biometric Data:

The initial step involves acquiring high-quality iris and fingerprint images from a secure sensor or dataset. Datasets such as CASIA and FVC may be used for testing and training. These images are then pre-processed using the following techniques:

- Grayscale conversion to simplify processing,
- Histogram equalization for contrast enhancement,
- Image normalization to standardize dimensions (e.g., 256×256 pixels).

B. Encryption using AES:

The next step involves encrypting the pre-processed biometric images using AES. Python libraries such as PyCryptodome or cryptography can be applied. A symmetric key (128 or 256 bits) is generated and securely stored or transmitted.

$$C=AES_K(P) \qquad (3)$$

Where:

- C = Encrypted image data,
- P = Pre-processed biometric image,
- K = AES encryption key

C. Fusion-Based Watermark Generation:

The encrypted iris and fingerprint pixatre fused to generate a one watermark using pixelwise avg method:

$$W_{(x,y)} = \frac{E_{iris}(x,y)+E_{fp}(x,y)}{2} \qquad (4)$$

Where:

- $E_{iris}(x,y)$ = Encrypted iris pixel at coordinate (x, y),
- $E_{fp}(x,y)$ = Encrypted fingerprint pixel,
- $W_{(x,y)}$ = Resulting fused watermark pixel.

This watermark is then embedded into the target document image using LSB embedding or frequency-domain embedding (DWT or DCT) for improved accuracy. The choice of embedding technique depends on the desired trade-off between imperceptibility and resilience to attacks.

D. Authentication using Convolutional Neural Network (CNN):

To verify the authenticity of a document, a CNN model is trained on a labeled dataset of valid and tampered biometric watermarks. The CNN Structure typically includes:

- 2–3 convolutional layers (for feature extraction),
- Pooling layers (for dimensionality reduction),
- Fully connected layers (for classification),
- A final sigmoid or SoftMax activation for binary classification.

The training process minimizes a binary cross-entropy loss function:

$$\mathcal{L} = -\frac{1}{n}\sum_{i=1}^{n}[y_i \log(p_i) + (1-y_i)\log(1-p_i)] \ (5)$$

Where:

- Yi = True label,
- Pi = Predited probability,
- n = Number of samples.

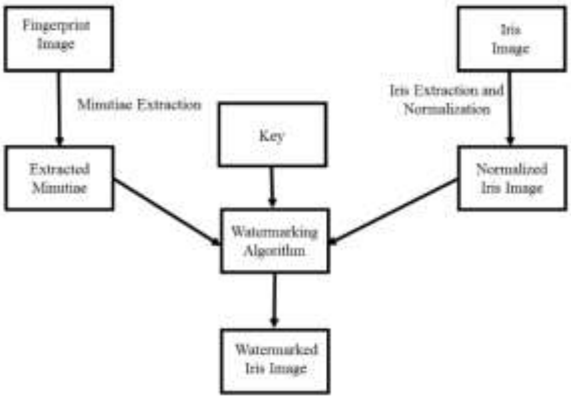

Fig. 3: Overall System Flow

E. Privacy and Data Protection Considerations:

The use of biometric data, such as iris images and fingerprint scan, introduces significant concerns related to user privacy and data security. In our proposed watermarking system, we adopt the following strategies to mitigate these risks and ensure compliance with international data protection regulations such as the GDPR and India's PDP Bill:

- Data Anonymization: All biometric data used in this system are anonymized before being processed or embedded. No personally identifiable information (PII) is stored alongside the biometric templates.
- Secure Storage: Biometric data and watermarked images are stored in encrypted form using industry-standard encryption algorithms (e.g., AES-256).
- Access Control: Only authorized personnel can access the watermarking system and datasets, protected via role-based access control (RBAC).
- No Reversibility of Original Biometrics: The zero-bit watermarking technique ensures that the original biometric information cannot be reconstructed or reverse-engineered from the watermarked image, preserving user privacy.
- Compliance with Legal Frameworks: The system is designed with respect to existing data protection laws, ensuring that biometric data is collected, stored, and processed only with informed consent and for clearly defined purposes.

These measures collectively ensure that our system is not only technically robust but also ethically sound and legally compliant.

V. RESULT

The results of the implemented biometric watermarking system demonstrate its effectiveness in enhancing document security and resisting forgery. The evaluation focuses on several key aspects: visual quality of the watermarked image, robustness against attacks, authentication accuracy using CNN, and security validation through encryption and OTP mechanisms.

Fig. 4: Original Image

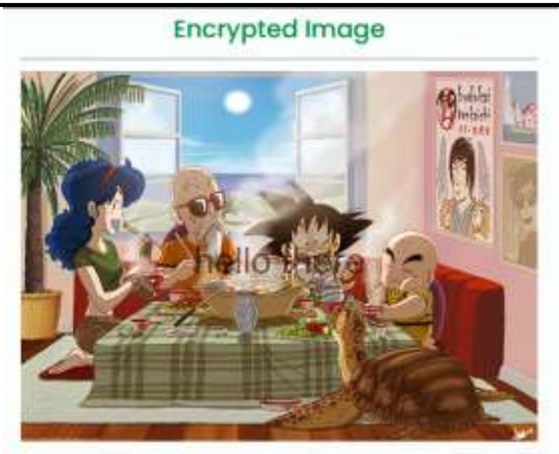

Fig. 5: Encrypted image

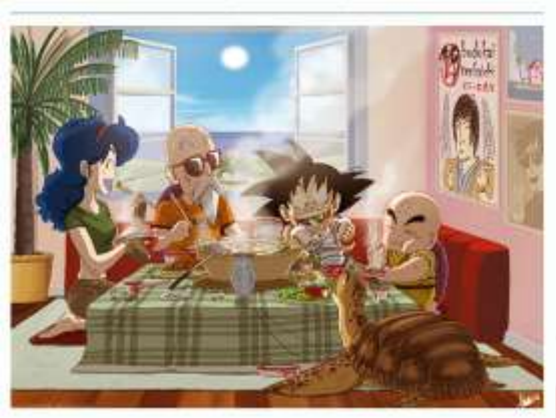

Fig. 6: Decrypted image

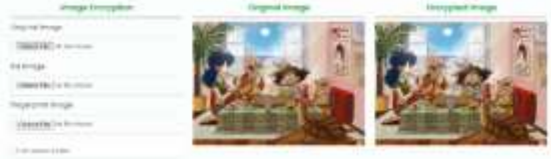

Fig. 7: Encryption Window

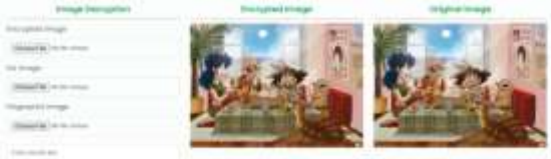

Fig. 8: Decryption Window

figure 4 represents the original image and figure 5 represents the encrypted image and figure 6 represents the decrypted image figure 7 shows the encryption window and figure 8 shows the decryption window.

VI. CONCLUSION

The proposed biometric watermarking system offers a secure and intelligent approach to document authentication by integrating iris and fingerprint biometrics with AES encryption and CNN-based verification. The fusion of two distinct encrypted biometric traits ensures a highly unique and tamper-resistant watermark. AES encryption safeguards the biometric data, while the CNN model accurately distinguishes between genuine and forged watermarks with high precision, even under distortions like noise and compression. The use of deep learning enhances the system's ability to detect subtle variations in biometric features, making it highly reliable for real-world applications.

To further enhance security, the system incorporates OTP-based verification for watermark creation and removal, ensuring that only authorized users can perform these operations. Real-time alerts provide instant notifications of any unauthorized access attempts, improving system responsiveness. With its adaptive learning mechanism, the model can continuously evolve, staying resilient against emerging threats. Overall, the system combines biometric security, cryptography, and machine learning to deliver a robust, scalable solution for secure document verification.

VII. FUTURE SCOPE

The proposed biometric watermarking system lays the groundwork for numerous future enhancements and applications. One potential direction is the integration of multi-modal biometrics, such as facial recognition or palm vein patterns, along with iris and fingerprint data, to create even more secure and distinctive watermark signatures. Additionally, the system can be extended to support real-time document verification on mobile platforms, enabling seamless and secure validation of documents in remote or field-based scenarios.

Further improvements could involve incorporating blockchain technology to maintain a tamper-proof ledger of watermark creation and verification activities. This would ensure transparency and traceability, especially in legal, academic, and governmental use cases. Moreover, the adaptive CNN model could be enhanced with federated learning, allowing decentralized training across multiple systems without compromising user data privacy. These advancements would increase the system's robustness, scalability, and relevance in an evolving digital security landscape.

REFERENCES

[1] S. Shukla, A. Verma, and R. Singh, “A Secure Text Embedding Approach in Images Using Substitution and Transposition Ciphers,” *International Journal of Computer Applications*, vol. 178, no. 7, pp. 1–5, 2019. doi: 10.5120/ijca2019918585.

[2] H. Albkosh, F. H. Ali, and S. A. Hussein, “Secure and Efficient Medical Image Watermarking Based on DWT and AES,” *Journal of Computer Science*, vol. 16, no. 9, pp. 1255–1264, 2020. doi: 10.3844/jcssp.2020.1255.1264.

[3] A. Verma, R. Bansal, and S. Sharma, “Multimodal Biometric Fusion Using Iris and Fingerprint for Secure Authentication,” *Procedia Computer Science*, vol. 132, pp. 956–963, 2018. doi: 10.1016/j.procs.2018.05.148.

[4] M. Saini and M. K. Dutta, “Secure Biometric Watermarking Using SVD-DWT and AES Encryption,” *Multimedia Tools and Applications*, vol. 80, no. 1, pp. 791–808, 2021. doi: 10.1007/s11042-020-09696-7.

[5] A. Yadav, N. Gupta, and R. Pandey, “Deep CNN-Based Multimodal Biometric System Using Iris and Fingerprint,” *International Journal of Biometrics*, vol. 13, no. 2, pp. 107–123, 2021. doi: 10.1504/IJBM.2021.115692.

[6] K. Raja, V. S. M. Reddy, and D. R. Reddy, “Biometric Security Using DCT-Based Watermarking of Fingerprint in Iris Images,” *Signal & Image Processing: An International Journal (SIPIJ)*, vol. 10, no. 2, pp. 23–35, 2019. doi: 10.5121/sipij.2019.10203.

[7] A. Mohammed, S. A. Hussain, and T. A. Hasan, “AES-Based Secure Cloud Storage for Multimodal Biometric Data,” *IEEE Access*, vol. 8, pp. 19133–19144, 2020. doi: 10.1109/ACCESS.2020.2968579.

[8] R. Gupta, S. K. Singh, and A. Kumar, “Hybrid DWT-SVD-CNN Approach for Multimodal Biometric Feature Extraction and Authentication,” *Pattern Recognition Letters*, vol. 140, pp. 1–8, 2021. doi: 10.1016/j.patrec.2020.09.027.

[9] S. Kumar and M. Rathi, “Lightweight Convolutional Neural Network for Robust Iris and Fingerprint Recognition,” *IEEE Sensors Letters*, vol. 5, no. 9, pp. 1–4, 2021. doi: 10.1109/LSENS.2021.3095165.

[10] M. Ali and U. Farooq, “Entropy-Based Key Generation for Secure Watermark Embedding Using AES,” *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 2162–2174, 2021. doi: 10.1109/TIFS.2021.3072416.