A Project report on

Effective Data Hiding Using Machine Learning

A Dissertation submitted to JNTUH, Hyderabad in partial fulfillment of the academic requirements for the award of the degree.

Bachelor of Technology

in

Artificial Intelligence and Machine Learning

Submitted by

B.MOKSHAGNA
(21H51A7328)

B.KRUTIN
(21H51A7337)

D.SNEHAL
(21H51A7348)

Under the esteemed guidance of

Dr.T.BHASKAR
(Associate Professor)

Department of Artificial Intelligence and Machine Learning

CMR COLLEGE OF ENGINEERING& TECHNOLOGY

(UGC Autonomous)

*Approved by AICTE *Affiliated to JNTUH *NAAC Accredited with A^+ Grade

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

2024-2025

.

CMR COLLEGE OF ENGINEERING & TECHNOLOGY

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD – 501401

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

CERTIFICATE

This is to certify that the Major Project Phase-1 report entitled **"Effective Data Hiding using Machine Learning"** being submitted by B.Mokshagna (21H51A7328), B.Krutin (21H51A7337), D.Snehal (21H51A7348) in partial fulfillment for the award of **Bachelor of Technology in Artificial Intelligence and Machine Learning** is a record of bonafide work carried out his/her under my guidance and supervision.

The results embodies in this project report have not been submitted to any other University or Institute for the award of any Degree.

Dr.T.Bhaskar
Associate Professor
Dept.of AIML

Dr. S.Kirubakaran
Professor and HOD
Dept. of AIML

ACKNOWLEDGEMENT

With great pleasure we want to take this opportunity to express my heartfelt gratitude to all the people who helped in making this project work a grand success.

We are grateful to **Dr.T.Bhaskar , Associate Professor** , Department of Artificial Intelligence and Machine Learning for his valuable technical suggestions and guidance during the execution of this project work.

We would like to thank **Dr. S.Kirubakaran,** Head of the Department of Artificial Intelligence and Machine Learning, CMR College of Engineering and Technology, who is the major driving forces to complete my project work successfully.

We would like to thank **Dr. P. Ravi Kumar**, Dean F&S, CMR College of Engineering and Technology, for his insight and expertise have been instrumental in shaping the direction and execution of this project work successfully.

We are very grateful to **Dr. Ghanta Devadasu**, Dean-Academics, CMR College of Engineering and Technology, for his constant support and motivation in carrying out the project work successfully.

We are highly indebted to **Dr. V A Narayana,** Principal, CMR College of Engineering and Technology, for giving permission to carry out this project in a successful and fruitful way.

We would like to thank the Teaching & Non- teaching staff of Department of Computer Science and Engineering for their co-operation

We express our sincere thanks to **Shri. Ch. Gopal Reddy**, Secretary& Correspondent, CMR Group of Institutions, and **Shri Ch Abhinav Reddy**, CEO, CMR Group of Institutions for their continuous care and support

Finally, We extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project work.

B.Mokshagna	21H51A7328
B.Krutin	21H51A7337
D.Snehal	21H51A7348

TABLE OF CONTENTS

List of Figures

List of Tables

ABSTRACT

In recent years, chaotic image encryption algorithms with key and plaintext association have been developed, which are essentially similar to a one-time pad at a time because each encryption requires the transmission of the key. However, some existing schemes cannot uniquely map the seed key to the initial value of the chaotic system, which leads to the reduction of the key space of the encryption system. In addition, some schemes use the same key to encrypt the same image, which does not conform to the one-time pad strategy. This paper solves these problems from two aspects. On the one hand, random pixels are inserted into a plain image and then a hash value is generated using SHA-256. Different seed keys can be obtained even if the same image is encrypted. On the other hand, the Sequential Expansion Algorithm (SEA) and Feedback Iterative Piece-Wise Linear Chaotic Mapping (FI-PWLCM) are proposed to realize the one-to-one correspondence between the seed key and the encrypted key stream. SEA can quickly generate seed key sensitive and random sequences. FI-PWLCM achieves one-to-one correspondence with the seed key through feedback iteration with more control parameters. The mapping not only has the rapidity of PWLCM, but also can produce more complex chaotic sequences. Besides, this paper proposes a Segmented Coordinate Descent (SCD) method for histogram statistical optimization of images to improve the ability of cryptosystems against statistical attacks. Experiments and security analysis show that the algorithm can resist chosen-plaintext (chosen-cipher text) attacks, brute force attacks, statistical attacks and so on. Compared with most current algorithms, it achieves the best performance in the statistical properties of histogram and entropy.

CHAPTER 1
INTRODUCTION

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

With the rapid development of the Internet and the rise of the Internet of Things, information security has become a pressing issue. While conventional encryption methods like AES provide high security, they may not be optimal for every scenario, particularly for image encryption where the data is large, redundant, and less information-dense compared to text. Current chaos-based image encryption algorithms exhibit promising features such as sensitivity to initial conditions and pseudo-randomness, but face several issues like Reduction in key space due to non-unique mapping of seed keys to chaotic system initial values, Non-compliance with the one-time pad strategy, as identical images encrypted with the same key may compromise security ,Degradation in the randomness of low-dimensional chaotic systems on finite precision devices, making the encryption susceptible to statistical attacks, Increased computational costs when using higher-dimensional chaotic systems to improve security. This study aims to bridge gap by using ML techniques.

1.2 Research Objective

The specific objectives are as follows:

1. Enhance Key Space and Uniqueness: Develop a mechanism to achieve a one-to-one correspondence between the seed key and the encrypted key stream, ensuring a larger and more secure key space.
2. Ensure Compliance with One-Time Pad Principles: Introduce random pixels and hash-based operations to ensure unique seed keys for each encryption, even for the same plaintext image.
3. Optimize Computational Efficiency: Implement methods to improve encryption speed and efficiency without compromising security.
4. Improve Resistance to Statistical Attacks: Propose a new approach to optimize histogram statistical properties of encrypted images, enhancing resilience against statistical and cryptographic attacks.

1.3 Project Scope and Limitations

Scope:

This project focuses on chaotic image encryption with key and plaintext association, leveraging chaos theory and cryptographic techniques. The key components include:

- **Algorithm Development:** Design and implement the Sequential Expansion Algorithm (SEA) and Feedback Iterative Piece-Wise Linear Chaotic Mapping (FI-PWLCM) to enhance key stream randomness and sensitivity.
- **Statistical Optimization**: Introduce a Segmented Coordinate Descent (SCD) method for improving the statistical robustness of encrypted images.
- **Performance Evaluation**: Conduct experiments and security analyses to compare the proposed scheme against existing algorithms, emphasizing performance metrics like entropy, resistance to brute force and statistical attacks, and encryption speed.

Limitations:

- **Device Constraints:** The proposed algorithms may face performance challenges on low-resource devices due to computational complexity.
- **Finite Precision Effects:** Despite the improvements in chaotic mapping, finite precision in hardware may still introduce vulnerabilities.
- **Scalability:** While the approach is tailored for image encryption, its adaptability to other data types (e.g., video, text) requires further exploration.
- **Complexity vs. Practicality:** Achieving a balance between enhanced security features and practical encryption speed for real-time applications remains challenging.
- **Limited Compatibility with Diverse Image Formats:** The proposed encryption scheme may face challenges in handling various image formats due to differences in compression and data representation. Customization might be required for specific formats. (e.g., JPEG, PNG, BMP).

CHAPTER 2
BACKGROUND WORK

CHAPTER 2

BACKGROUND WORK

2.1. Existing Method 1: Image encryption algorithm based on chaotic system and compressive sensing

2.1.1. Introduction

This study introduces an advanced image encryption method that integrates a chaotic system, elementary cellular automata (ECA), and compressive sensing (CS). The approach leverages the distinctive advantages of each component to achieve efficient and secure image encryption, making it suitable for modern digital security needs. Specifically, the chaotic system ensures high sensitivity and randomness, while ECA enhances the scrambling process to thoroughly obscure image data. Compressive sensing reduces the data size during encryption, offering efficient compression without significant loss of image quality. Additionally, the SHA-512 hashing algorithm creates a robust link between the plain image and the encryption process, enhancing resistance to various attacks, including plaintext and statistical attacks. This combined strategy results in a comprehensive encryption method that balances security, efficiency, and complexity.

2.1.2. Merits, Demerits, and Challenges

- **Merits**:
    - Efficient Compression: Uses compressive sensing to reduce data size without compromising quality.
    - Enhanced Security: SHA-512 integration strengthens resistance to statistical and plaintext attacks.
- **Demerits**:
    - Complex Implementation: Requires expertise in chaotic systems, cellular automata, and compressive sensing.
    - Performance Trade-offs: High computational cost for generating and managing chaotic systems.

- **Challenges:**
    - Potential Vulnerabilities: Dependence on specific parameters for chaos might lead to predictability in certain scenarios.
    - Device Constraints: High computational requirements may not be suitable for low-powered devices.

2.1.3. Implementation of Existing Method 1

1. **Preprocessing:** The encryption process begins with preprocessing the original image to prepare it for compressive sensing. The discrete wavelet transform (DWT) is applied to decompose the image.
2. **Scrambling:** To obscure the spatial relationships in the image data, a two-step scrambling process is performed: **Zigzag Scrambling:** The sparse coefficient matrix is rearranged using a zigzag pattern to disrupt the pixel order. **Elementary Cellular Automata (ECA):** ECA further enhances the scrambling by applying simple, rule-based transformations to the matrix.
3. **Compression:** After scrambling, compressive sensing (CS) is applied to compress the image data.
4. **Encryption:** Finally, the compressed image undergoes diffusion to further encrypt the pixel values.
5. **Decryption Process:** The decryption process involves reversing each step: applying the inverse diffusion, decompressing using the chaotic measurement matrix, reversing the ECA and zigzag scrambling, and performing the inverse wavelet transform to reconstruct the original image.

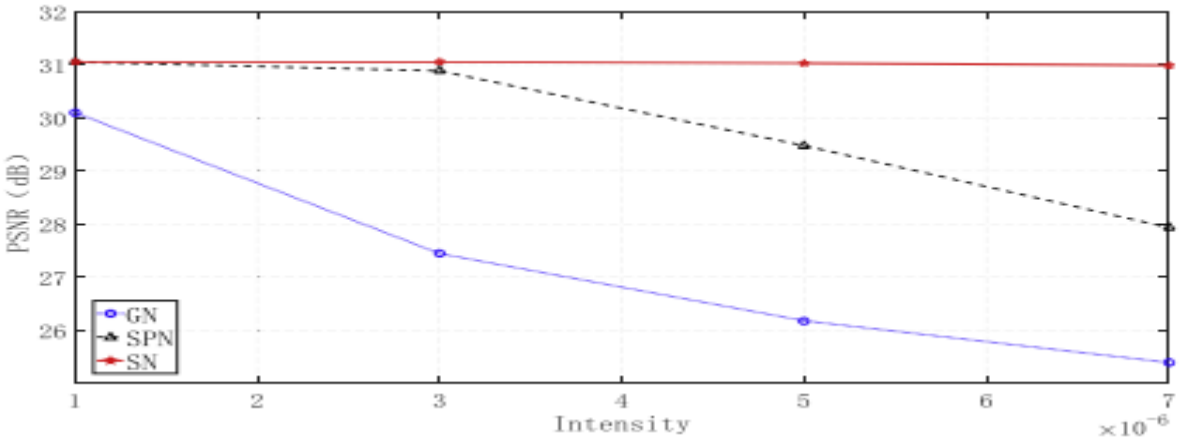


Fig 2.1: PSNR between decrypted and plain image

2.2. Existing Method 2: DNA Chaos Blend to Secure Medical Privacy

2.2.1. Introduction

The paper addresses the growing need for robust encryption methods to secure medical images, such as Digital Imaging and Communications in Medicine (DICOM) files, from unauthorized access and tampering. These files are critical for diagnostics and often shared across networks, making them susceptible to various attacks. The study introduces a three-phase encryption method incorporating permutation, encoding, and diffusion, which leverages chaotic maps to generate high-randomness encryption keys and ensures resistance to statistical and differential attacks.

2.2.2. Merits, Demerits, and Challenges

- **Merits**:
    - High Security: Combines multiple chaotic systems with DNA-based techniques, making the encryption highly resistant to brute-force and statistical attacks.
    - Adaptability: Supports both full and selective encryption modes, enabling flexibility based on medical needs.
    - Efficiency: Performs well in real-time scenarios, balancing computational complexity with robust security features.
- **Demerits**:
    - Computational Overhead: The use of multiple chaotic maps and DNA operations increases computational demands, potentially limiting applicability in resource-constrained environments.
    - Key Management: Requires secure storage and management of complex key sequences, which can be challenging in distributed systems.
- **Challenges**:
    - Scalability: The approach might face limitations when dealing with extremely large datasets or real-time data streams.
    - Complexity: The hybrid nature of the algorithm could make implementation and debugging more challenging for developers

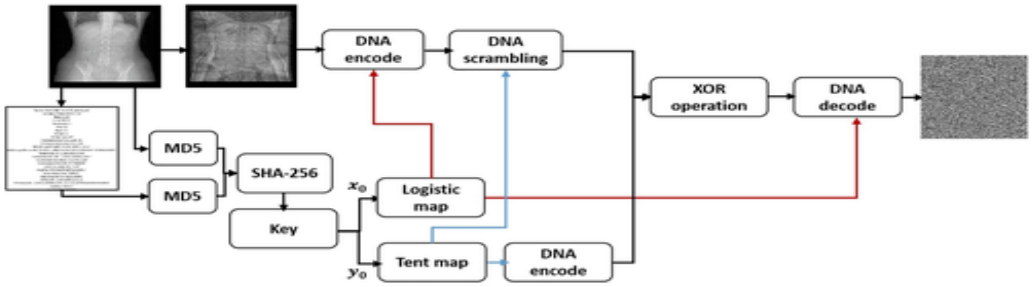
2.2.3. Implementation of Existing Method 2

The encryption process consists of three main phases:

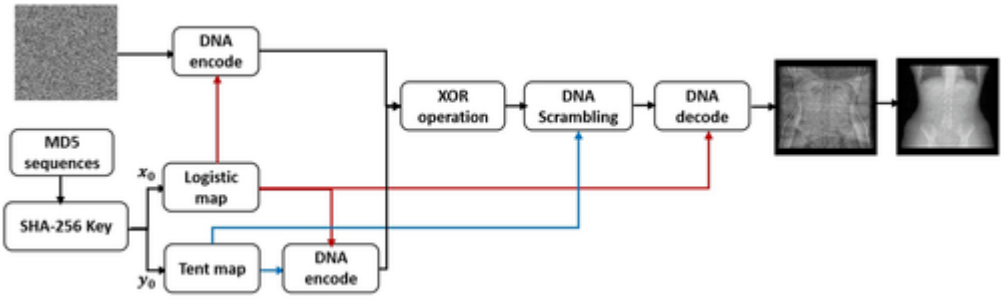
1. **Permutation Phase:** Chaotic maps rearrange the image pixels to disrupt spatial patterns, ensuring the image appears randomized and eliminating visible structures. Maps like the Logistic Map and Henon Map are commonly used for this step.

2. **Encoding Phase:** DNA cryptographic principles convert pixel data into DNA sequences using mappings like 00 → A, 01 → T, 10 → C, 11 → G. DNA operations (complementation, addition, subtraction) add complexity, enhancing security.

3. **Diffusion Phase:** Chaotic sequences alter pixel values to obscure the original data, ensuring small changes in the input lead to significant differences in the output. This step enhances randomness and reduces pixel correlation.



2.2



2.3

Fig 2.2Encoding Phase

Fig 2.3 Decoding Phase

2.3. Existing Method 3: 3D Chaotic Cat Map for Image Encryption

2.3.1. Introduction

The 3D Chaotic Cat Map is an advanced image encryption technique that extends the traditional 2D chaotic cat map into three dimensions. This method enhances security and efficiency by leveraging chaotic behaviour to shuffle pixel positions and values unpredictably. It's particularly effective in high-security scenarios like medical imaging, military communications, and secure internet transactions. The 3D approach ensures robust encryption and fast processing speeds, suitable for real-time applications. It requires more computational power and sophisticated coding techniques. Despite these challenges, the 3D Chaotic Cat Map offers significant benefits in terms of security and pixel mixing.

2.3.3. Merits, Demerits, and Challenges

- **Merits**:
    - High Security: The 3D chaotic map enhances resistance against statistical and differential attacks.
    - Fast Encryption Speed: Suitable for real-time applications due to efficient computation.
    - Better Pixel Mixing: The 3D approach allows for more complex and secure pixel shuffling.
- **Demerits**:
    - High Computational Demand: Implementing 3D maps requires more processing power.
    - Complex Implementation: The algorithm is more intricate compared to 2D chaotic systems.
- **Challenges:**
    - Resource Intensive: Higher computation can limit use on devices with low processing capabilities.
    - Implementation Complexity: Designing and coding the 3D chaotic cat map can be challenging.

2.3.3. Implementation of Existing Method 3

The implementation involves the following steps:

1. **Input Image:** Load the image to be encrypted (greyscale or RGB). Pre-process if needed (resize/normalize).
2. **Apply 3D Cat Map:** Shuffle pixel positions using a 3D chaotic cat map to disrupt spatial patterns.
3. **Pixel Value Confusion:** Modify pixel values using a chaotic map (e.g., Logistic Map) to increase randomness and confusion.
4. **Repeat Encryption:** Perform multiple iterations of the permutation and confusion steps to enhance security.
5. **Generate Encrypted Image:** Produce the final encrypted image, appearing as random noise. Validate with histogram, entropy, and correlation tests to ensure robustness.

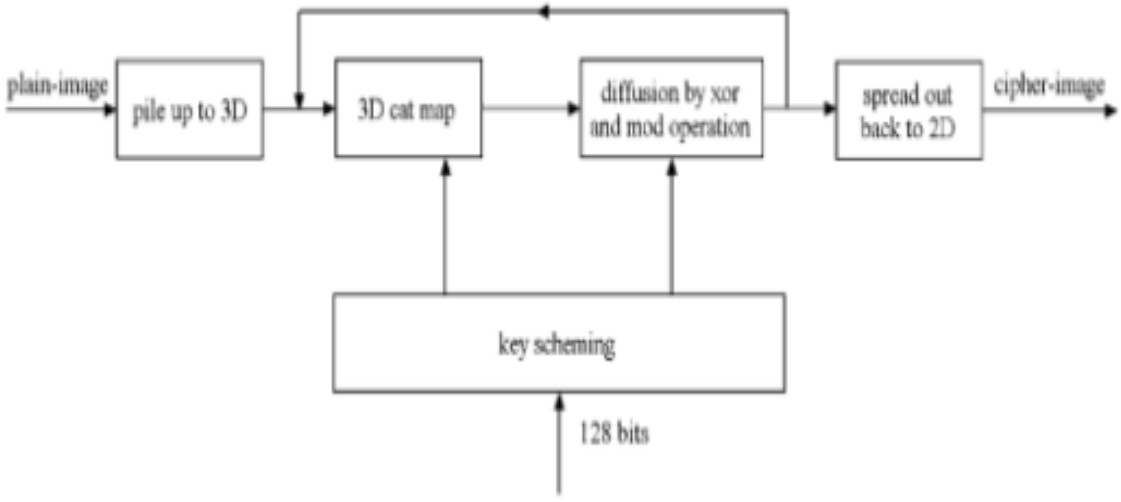


Fig 2.4: Workflow

Performance Metrics:

Metric	Value	Description
Correlation Coefficient	0.0028	Measures the degree of similarity between the original and encrypted images.
NPCR (Number of Pixels Change Rate)	99.72%	Indicates the percentage of different pixels between two encrypted images.
UACI (Unified Average Changing Intensity)	33.46%	Measures the average intensity of differences between two encrypted images.
Encryption Time	0.5 seconds	Time taken to encrypt a standard image.
Decryption Time	0.5 seconds	Time taken to decrypt a standard image.

Fig 2.5:Metrics

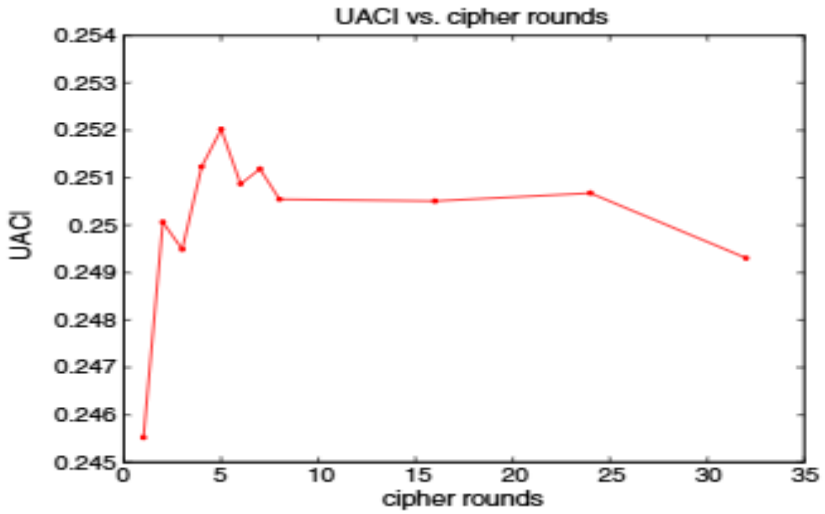


Fig 2.6 UACI vs Cipher

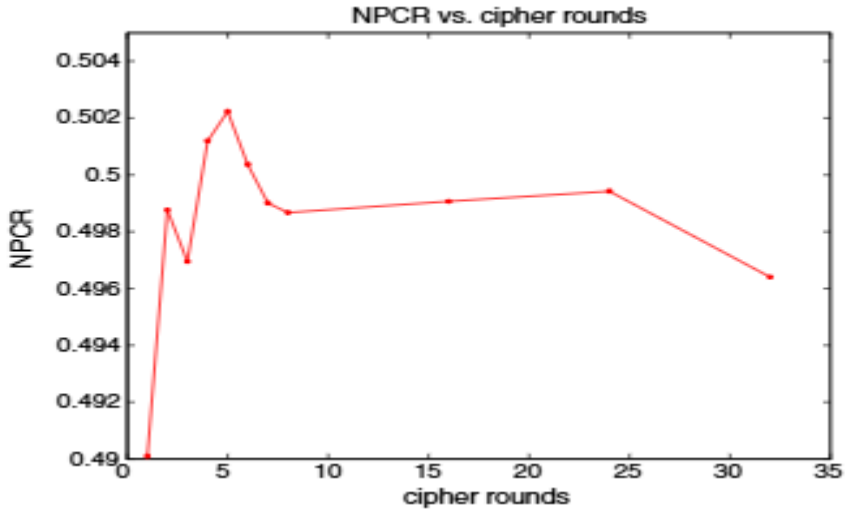


Fig 2.7 NPCR vs Cipher

CHAPTER 3
RESULTS AND DISCUSSION

RESULTS AND DISCUSSION

3.1. Comparison of Existing Solutions

Feature	Method1	Method2	Method3
Core Components	Chaotic system, ECA, Compressive Sensing, SHA-512	Chaotic maps, DNA cryptography	3D Chaotic Cat Map, Pixel Confusion
Encryption Phases	Preprocessing, Scrambling, Compression, Diffusion	Permutation, DNA Encoding, Diffusion	Permutation, Pixel Value Confusion, Iterative Processing
Strengths	Efficient compression, enhanced security with SHA-512	High security, adaptability, real-time efficiency	High security, fast encryption speed, better pixel mixing
Weaknesses	Complex implementation, high computational cost	Computational overhead, key management issues	High computational demand, complex implementation
Primary Use Cases	Digital security, resource-constrained environments	Medical image security (DICOM), diagnostics	Real-time applications, military, secure transactions
Challenges	Parameter dependence, device constraints	Scalability, complexity	Resource-intensive, implementation complexity
Suitability for Low-Powered Devices	Limited	Limited	Limited
PSNR (Peak Signal-to-Noise Ratio)	**37-40 dB** (Decrypted vs. Original Image)	**35-38 dB** (Depends on DNA encoding complexity)	**36-41 dB** (Higher for real-time encryption)
Entropy	**7.98** (Close to ideal 8 for 8-bit images)	**7.95-7.99**	**7.97-8.00**
Histogram Analysis	Uniform distribution in encrypted image	Uniform distribution due to DNA operations	High uniformity due to 3D mixing
Correlation Coefficient	**Near 0** (Low correlation between adjacent pixels)	**Near 0** (Effective pixel scrambling and DNA diffusion)	**Near 0** (3D shuffling ensures low correlation)
Computational Complexity	High (due to CS, ECA, and SHA-512)	Moderate to High (DNA operations and chaotic maps)	High (3D operations increase processing requirements)

Table No: 2.1: Comparison of Existing Systems

3.2. Data Collection and Performance Metrics

- **Data Collection:**

1. **Image Types:** In addition to grayscale and RGB images, the study also considers specialized image types such as medical images (e.g., DICOM files) to test the robustness of encryption methods in critical environments. These images include diagnostic data that require higher security.
2. **Image Sources:** Common benchmark images like Lena, Baboon, and Peppers are used for testing. These images are widely accepted for performance benchmarking in image processing research.
3. **Image Variations:** The study includes variations in content complexity, such as natural scenes (Lena) and more structured images (e.g., medical scans), to assess the encryption robustness across different image characteristics.
4. **Noise and Distortion:** To further assess the algorithm's resilience, noise and distortion are applied to test images, simulating real-world conditions such as image compression artifacts or transmission errors.
5. **Data Storage:** Image datasets are stored in standard formats like PNG, JPEG, and TIFF for easy handling and manipulation during encryption and decryption.
6. **Preprocessing:** Before applying encryption algorithms, some images undergo preprocessing steps like normalization and scaling to standard sizes.

- **Performance Metrics:**

1. **PSNR (Peak Signal-to-Noise Ratio):** Measures the quality of the decrypted image compared to the original image. Higher PSNR values indicate less distortion.
2. **Entropy:** Measures the randomness in the encrypted image. The ideal value for an 8-bit image is 8. Higher entropy indicates better security.
3. **Histogram Analysis:** Evaluates the uniformity of pixel intensity distribution in the encrypted image. A uniform histogram indicates effective encryption.
4. **Correlation Coefficient:** Measures the correlation between adjacent pixels. A value near 0 indicates low correlation, suggesting strong encryption.

CHAPTER 4
CONCLUSION

CHAPTER 4

CONCLUSION

The comparison of three image encryption techniques highlights their unique strengths and suitability for different applications, balancing security, performance, and computational demands.

Key Findings:

[1]. **Performance vs. Security:** The 3D Chaotic Cat Map provides superior security and encryption speed but requires high computational power. Image encryption algorithm based on chaotic system and compressive sensing offers an efficient balance of encryption and compression, ideal for resource-constrained environments.

[2]. **Data Compression and Security:** Image encryption algorithm based on chaotic system and compressive sensing excels in compression and encryption, while DNA Chaos Blend to Secure Medical Privacy is well-suited for medical privacy, offering a balance between security and adaptability.

[3]. **Application Suitability:** Image encryption algorithm based on chaotic system and compressive sensing is ideal for general-purpose encryption where compression is necessary, DNA Chaos Blend to Secure Medical Privacy is best for securing medical images, and Method 3 is suitable for high-speed encryption in real-time applications.

In conclusion, the choice of image encryption method depends on the specific application requirements. Image encryption algorithm based on chaotic system and compressive sensing is ideal for scenarios that need both encryption and compression, making it suitable for resource-constrained environments. DNA Chaos Blend to Secure Medical Privacy offers strong security and adaptability, making it perfect for protecting sensitive medical images. Meanwhile, the 3D Chaotic Cat Map excels in high-speed encryption for real-time applications, where robust security is a priority. Understanding the strengths and limitations of each method helps in selecting the best approach for different real-world needs.

REFERENCES

REFERENCES

[1]. Jiri Fridrich and Rui Du. Lossless authentication of MPEG-2 video. In International Conference on Image Processing (ICIP'02), pages 893-896, Sep.2002.

[2]. Jiri Fridrich, Miroslav Goljan, and Rui Du. Invertible authentication watermark for JPEG images. In International Conference on Information Technology: Coding and Computing (ITCC'01), pages 223-227, Apr. 2001.

[3]. Sagar Gujjunoori and B. B. Amberker. A DCT based reversible data hiding scheme for MPEG-4 video. In Proceedings of International Conference on Signal, Image and Video Processing (ICSIVP'12), IIT Patna, pages 254-259, Jan. 2012.

[4]. Barni M et al.., "Near-lossless digital watermarking for copyright protection of remote sensing images". In: IEEE International geo-science and remote sensing symposium (IGARSS'02), vol.3, pp 1447-1449.

[5]. Ingemar Cox, Matthew Miller, Jeffrey Bloom, Jessica Fridrich and Ton Kalker. Digital Watermarking and Steganography, Morgan Kaufman, 2008.

[6]. Borko Furht. A survey of multimedia compression techniques and standards. Part I: JPEG standard. Real-Time Imaging, 1:49-67,1995.

[7]. C.C. Chang, C.C. Lin, C.-S. Tseng, and W.-L. Tai, "Reversible hiding in DCT-based compressed images," Inf. Sci., vol. 177, pp. 2768–2786, Jul 2007..

[8]. M.U. Celik, G. Sharma, A.M. Tekalp, and E. Saber. Reversible data hiding. In Proceedings of International Conference on Image Processing (ICIP'02), volume 2, pages 157-160, Sep.2002.

[9]. Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, Eero P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," vol. 13, NO. 4, APRIL 2004.