1 Introduction

DNA data storage is an existing technology that, while still experimental, is rapidly evolving. The ability to store information in synthetic DNA exploits the high storage density that DNA offers, as one gram of DNA can theoretically hold approximately 215 petabytes of data[2] . This makes DNA storage extremely attractive for storing large volumes of data in a very small physical space [3, 4]. However, there is a critical need that makes the study of this technology even more urgent and appealing: the storage capacity currently used for digital data will not suffice in the future. Forecasts suggest that traditional storage media, such as hard drives and cloud systems, will reach their limits within the next few decades [5, 6]. Therefore, the need for new storage methods, such as DNA storage, is becoming increasingly pressing.

At the same time, the synthesis of synthetic DNA has become feasible and has advanced significantly in recent years. Developments in biotechnology now allow the construction of synthetic DNA with high precision, and commercial companies, such as IDT [7], GenScript [8], and Twist Bioscience [9], offering custom gene synthesis services [10]. This means that the encrypted sequences created can be stored in physical DNA molecules and retrieved through sequencing techniques, thus enhancing the potential use of DNA for cryptographic purposes.

In this context, encryption through DNA computing leverages the DNA storage infrastructure to store encrypted information in natural or synthetic DNA molecules [11,12]. This storage approach is combined with encryption methods based on chaotic systems (such as the logistic map), which have been widely used to generate unpredictable cryptographic keys. Using such methods further strengthens the security and resilience of the encrypted data.

Moreover, embedded systems like the ESP32 are used to collect real-time environmental data (e.g., temperature, humidity, pressure), dynamically modifying the parameters of the chaotic system. This makes the encryption process dynamic and unpredictable [13]. Such technologies are increasingly utilized in modern security applications, enhancing the adaptability of cryptographic systems to diverse environments and threats.

1.1 Aims and Objectives

The objective of this thesis is to design, develop, and implement a novel hybrid encryption framework that combines advanced cryptographic, biological, and chaotic system techniques to maximize the confidentiality, integrity, and resilience of sensitive data. Specifically, the research aims to:

• Utilize Huffman encoding to enhance data entropy and minimize redundancy, disrupting statistical patterns that could be exploited in cryptanalysis.

• Encode and encrypt information through DNA computing, leveraging DNA's massive storage density and unique representational structure to design biologically plausible encrypted sequences suitable for synthetic DNA storage

• Apply biological modifications to create encrypted DNA sequences that closely resemble natural genomic material, thereby increasing obfuscation and making unauthorized decryption substantially more difficult.

• Incorporate chaotic systems based on logistic maps, dynamically influenced by real-time environmental data gathered from IoT devices like the ESP32, to generate encryption keys that are highly unpredictable and environment-specific, strengthening resistance against attacks based on key prediction or static assumptions.

• Implement iterative and fractional chaotic encryption mechanisms, introducing multiple layers of dynamic transformation that exponentially increase the system's cryptographic complexity while preserving the full reversibility needed for accurate decryption.

Through this interdisciplinary approach, the proposed system aspires to set a new standard for multi-layered, chaos-enhanced encryption, tailored specifically for fields requiring extreme data protection, such as medical imaging archives, genomic databases, intellectual property safeguarding, and critical infrastructure record-keeping. The focus is placed on security rather than processing speed, recognizing the increasing need for encryption systems that prioritize robustness and future-proofness in an era of escalating cyber threats.

1.2 Research Questions

1. How do simulated biological features (e.g., CRISPR, introns) increase structural complexity and disrupt statistical regularity, as demonstrated by Chi-Square and NIST STS results?
2. How does chaotic key modulation based on real-time ESP32 sensor data contribute to non-reproducible key generation and improved encryption randomness?
3. How does the full encryption–decryption pipeline preserve perfect reversibility, achieving SSIM = 1.0 and PSNR = ∞ in image reconstructions?
4. How does the proposed multi-layer encryption framework scale in time and memory with file size, and what trade-offs arise in real-world applications?
5. How feasible is it to store the final encrypted DNA outputs in synthetic DNA, in terms of base length and physical DNA mass?

1.3 Contribution of This Work

This thesis contributes a novel hybrid encryption framework that combines Huffman encoding, AES, DNA-based cryptography with biological transformations, chaos-based key generation, and real-time embedded sensor integration. Specifically, it:

- Proposes a new use of Huffman coding not just for compression but also to improve diffusion and entropy in the early stages of encryption.
- Introduces biological modifications—such as simulated intron/exon manipulation and CRISPR logic—to enhance complexity and resist statistical attacks in DNA cryptographic layers.
- Utilizes chaos theory, particularly logistic maps, for dynamic key generation driven by real-time environmental data.
- Implements and evaluates the ESP32 embedded platform for generating non-reproducible cryptographic keys from sensor inputs, validating practical feasibility in constrained environments.

- Develops a layered encryption-decryption mechanism that ensures reversibility while maximizing unpredictability and security.
- Explores storage of encrypted data within synthetic DNA structures, demonstrating potential for scalable and long-term data security.
- Offers a comprehensive analysis of security, entropy, and computational performance across all layers, with a focus on medical imaging as a practical case study.

By addressing underexplored intersections between biological, chaotic, and embedded technologies, this thesis provides a blueprint for secure, lightweight, and biologically integrated cryptographic systems suitable for both modern and future challenges.

1.4 Structure of the Thesis

This thesis is organized into six main chapters:

- **Chapter 1: Introduction**
  Presents the motivation, research questions, objectives, and contributions of this work, setting the context for the hybrid encryption system proposed.
- **Chapter 2: Theoretical Background**
  Provides a comprehensive overview of the foundational concepts, including classical and modern cryptographic techniques, Huffman coding, DNA computing, chaos theory, and embedded systems. It also introduces biological transformations and storage mechanisms in synthetic DNA.
- **Chapter 3: Related Work**
  Surveys the current literature across domains relevant to this research, categorizing prior work in DNA cryptography, chaos-based systems, biological obfuscation, DNA storage, and embedded platforms. It also identifies existing research gaps and presents the research questions.
- **Chapter 4: Methodology and System Implementation**
  Details the design of the proposed hybrid cryptographic system, including the integration of Huffman encoding, DNA logic, chaotic key generation, and embedded real-time data. It also covers the encryption and decryption processes, data flow, and platform-specific implementations.
- **Chapter 5: Result Extraction and Evaluation Methodology**
  Detailed analysis of the functions used within the code to assemble and evaluate the results of the proposed methodology.
- **Chapter 6: Evaluation and Results**
  Presents the experimental setup, security analysis, performance benchmarks, and scalability testing. The results are compared with existing schemes to highlight the advantages and trade-offs of the proposed approach.
- **Chapter 7: Conclusion and Future Work**
  Summarizes the findings of the research, reflects on limitations, and outlines directions for further development and real-world application of biologically integrated hybrid cryptographic systems.

1.5 Summary

This chapter introduced the motivation for developing a biologically inspired hybrid encryption framework that integrates Huffman encoding, DNA computing with biological obfuscation, chaotic systems, and embedded sensor-based key generation. It outlined the aims and objectives of the research, formulated key research questions, and defined the novel contributions made by this thesis. The proposed system targets high-security applications such as medical imaging and archival DNA storage, emphasizing robustness, complexity, and future-proof adaptability. The chapter concluded with an overview of the thesis structure, which guides the reader through the theoretical foundations, the literature context, the methodological implementation, the evaluation metrics, and the final conclusions. The next chapter provides a comprehensive background on the cryptographic, biological, and computational concepts essential for understanding the proposed framework.

2 Theoretical Background

2.1 Cryptography Background

2.1.1 Classical Methods

Classical cryptographic methods refer to encryption techniques that were developed and used before the rise of modern, computer-based cryptography (roughly before the 1970s). These methods usually rely on simple mathematical operations like substitution and transposition, rather than complex algorithms based on number theory, group theory, or elliptic curves [14].

2.1.1.1 Caesar Cipher

Caesar's Cipher is one of the oldest and simplest Classical Cryptographic Methods. It belongs to the substitution ciphers. Each letter in the plaintext is shifted a certain amount of positions down or up the alphabet [14]. Suppose we choose a shift of 3.

Plaintext	Ciphertext	Plaintext	Ciphertext	Plaintext	Ciphertext
A	D	J	M	S	V
B	E	K	N	T	W
C	F	L	O	U	X
D	G	M	P	V	Y
E	H	N	Q	W	Z
F	I	O	R	X	A
G	J	P	S	Y	B
H	K	Q	T	Z	C
I	L	R	U		

Table 2.1: Caesar Cipher Letter Mapping with a Shift of 3

The Caesar cipher, when implemented without supplementary encryption mechanisms, fails to meet contemporary security standards due to its limited key space and statistical predictability.

2.1.1.2 Vigenère Cipher

This cipher is an improvement over the Caesar one, as it is a polyalphabetic substitution cipher, using whole words/keys to encrypt a message. Each letter of the key tells you how much to shift each letter of the message [14, 15].

How it works:

1. Choose a key word.

Example: KEY

2. Repeat the key until it matches the length of your message.

Example message: ATTACKATDAWN Key repeated: KEYKEYKEYKEY

3. Encrypt by shifting each letter of the message by the amount corresponding to the matching key letter:

Plaintext Letter	Key Letter	Shift (Key Position - 1)	Ciphertext Letter
A	K	10	K
T	E	4	X
T	Y	24	R
A	K	10	K
C	E	4	G
K	Y	24	I
A	K	10	K
T	E	4	X
D	Y	24	B
A	K	10	K
W	E	4	A
N	Y	24	L

Table 2.2: Vigenère Cipher Encryption Example using the Keyword KEY

Final encrypted text: KXRKGIKXBKAL

• If the key is short or repeats too much, patterns still leak.

• If the key is as long as the message and completely random, it becomes theoretically unbreakable.

2.1.2 Symmetric Cryptographic Methods

Symmetric Cryptographic Methods(also called Symmetric Cryptography or Symmetric-key Encryption), are Cryptographic Techniques, where the same key is used for both encrypting and decrypting the data [16].

Main Characteristics:

• Speed: Symmetric encryption is very fast and efficient, especially for large amounts of data.

• Security: Security depends heavily on keeping the secret key private. If someone else gets the key, they can easily decrypt everything.

• Key management challenge: The key must be shared safely between the sender and the receiver. If you cannot securely share the key, the whole system is at risk.

2.1.2.1 AES – Advanced Encryption Standard

The Advanced Encryption Standard (AES) is a symmetric block cipher adopted by NIST in 2001 as the successor to DES [16]. It operates on 128-bit data blocks and supports key lengths of 128, 192, or 256 bits. The number of transformation rounds, 10, 12, or 14, depends on the key size. Each round includes the operations: SubBytes (S-box substitution), ShiftRows (bytewise row shift), MixColumns (column mixing) and AddRoundKey (XOR with key schedule) [16].