University of Zawia
Faculty of Science
Department of Computer Science

Preserving the Integrity and Authenticity of DICOM Medical Image Through Reversible Watermarking with Integer Wavelet Transform and adaptive Prediction Error Expansion

Presented by the student: Asma Khalifa Alnaef

Under supervision of: Dr. Nuha Omran Abokhdair
(Associate Professor)

Thesis Was Submitted in Partial Fulfilment of the Requirements for The Degree of Master in Computer Science.

Date (11/2/2026)

Declaration

Declaration I Asma Khalifa Alnaef confirm that the work contained in this thesis / dissertation, unless otherwise referenced is the researcher's own work, and has not been previously submitted to meet requirements of an award at this University or any other higher education or research institution, I furthermore, cede copyright of this thesis / dissertation in favour of University of Zawia.

Student name: Asma Khalifa Alnaef

Signature: .....

Date: ..... / ...../ 20

Dedication

To the soul of my beloved father...

My God have mercy on him and forgive him.

To my dear mother...

My God grant her health and well-being.

To my dear husband,

partner of life and companion on this journey.

To the delights of my heart, joy of my life...

my children Omar, Yusuf, Jana, and Islam.

To my dear family, my husband's family, and all my relatives' ones

who were my strength and support at every step.

To my esteemed professors

who illuminated my path with knowledge and wisdom,

To all friends,

I dedicate this humble work to you.

Asma Khalifa Alnife

Acknowledgement

I extend my deepest gratitude and appreciation to Almighty God, who guided me and facilitated the completion of this thesis.

I express profound thanks to my academic supervisor, Dr. Nuha Omran, for her dedicated effort and time in guiding and supporting me throughout this academic journey, which significantly contributed to developing my research vision.

I also thank the distinguished members of the thesis committee for their valuable comments and contributions in refining this work.

I extend special thanks to my professors in the Department of Computer Science at the University of the Zawia, who generously shared their knowledge and expertise with me.

I would also like to express my sincere appreciation to the College of Computer Technology in Zawia for granting me study leave, which enabled me to dedicate myself fully to this research.

I express my deep appreciation to the University of the Zawia for providing me with the opportunity to conduct research and study.

Finally, I offer sincere thanks to my family for their patience and love, which served as my greatest motivation throughout this academic journey.

Researcher

List of Contents

List of Tables

List of illustrations

List of Appendices

List of abbreviations

Abbreviation	Meaning
APEE	Adaptive Prediction Error Expansion
BER	Bit Error Rate
CT	Computed Tomography
DB	Decibel
DCT	Discrete Cosine Transform
DE	Difference Expansion
DICOM	Digital Imaging and Communications in Medicine
DWT	Discrete Wavelet Transform
EPR	Electronic Patient Records
HIPAA	Health Insurance Portability and Accountability Act
IWT	Integer Wavelet Transform
LWT	Lifting Wavelet Transform
MRI	Magnetic Resonance Imaging
NCC	Normalized Cross Correlation
PEE	Prediction Error Expansion
PSNR	Peak Signal to Noise Ratio
ROI	Region of Interest
RONI	Region of Non-Interest
SHA	Secure Hash Algorithm
SSIM	Structural Similarity

ABSTRACT

The widespread adoption of telemedicine and medical image exchange over public networks has raised concerns about the security and integrity of sensitive patient data. Digital watermarking has emerged as a promising technique for securing exchanged medical images and hiding the information of patients, providing confidentiality, authentication, and tamper detection. Reversible watermarking methods are particularly valuable for medical applications, as they enable the complete recovery of the original image after watermark extraction, preserving diagnostic quality. To enhance robustness against attacks, achieve high capacity, and ensure imperceptibility, this research introduces a novel reversible watermarking scheme for brain MRI and CT images. that integrates the lifting-based integer wavelet transform with an adaptive prediction error expansion is introduced in order to ensure the integrity, authenticity, and confidentiality of the medical images and patient information. To enhance embedding efficiency, the scheme segments the coefficients of the medical image using an unsupervised machine learning technique into Region of Interest (ROI) and Region of Non-Interest (RONI). To preserve the quality of the diagnostically significant region, one bit is hidden in each embeddable coefficient in ROI, while two bits are hidden in each embeddable coefficient in RONI. The hash value of the ROI and patient data are embedded in the ROI, while the tamper detection and recovery and the hash value of the RONI are embedded in the RONI. The robustness of the proposed scheme was evaluated against Gaussian noise attacks under various intensity levels and intended attacks, including cropping, shearing, and collage attacks. The experimental results validate the proposed scheme's efficiency, reversibility, data integrity verification, privacy and security protection, and high embedding capacity with high image quality. Additionally, it accurately detects and recovers tampered regions with high quality

Chapter 1

Introduction

1.1 Introduction

Recent rapid development in information and communication technologies has enhanced healthcare services, enabling practices such as telemedicine. This enables the exchange of medical images and electronic health records between doctors, patients, and hospitals. However, communication networks expose these images to interception and tampering by intruders. [1, 2]. Despite the protective measures, millions of security breaches involve medical data every year, with 15% of all cyberattacks globally detected in a healthcare[3, 4].

To ensure medical image integrity and patient privacy, regulations like the Health Insurance Portability and Accountability Act (HIPAA) mandate specific security standards[5].The Digital Imaging and Communications in Medicine (DICOM) standard addresses this through authentication, confidentiality, and integrity profiles[6].

The secure transmission of medical images depends on these properties. Confidentiality ensures data protection from unauthorized access, allowing only authorized users to access information. Integrity ensures that the data sent matches the data received without any alterations. Authentication verifies sender identity and confirms that the data pertains to the correct patient [7, 8].

Various security techniques have been developed to address these challenges, where digital watermarking and encryption dominate telemedicine security [2, 9], Encryption obscures data, while watermarking embeds ownership or authentication markers to verify origin and integrity [9]. However, limitations of DICOM persist: digital signatures protect pixel data integrity but not confidentiality, and privacy profiles secure headers without ensuring their authenticity and integrity. Post-decryption integrity checks falter if signatures are removed, and signatures require periodic renewal [10, 11]. Among these security techniques, digital image watermarking has garnered significant attention from researchers; due to its unique characteristics that make it well-suited for resolving the challenges of safeguarding the copyright and security of sensitive medical information and patient diagnostic data [12, 13]. Moreover, Watermarking offers more detailed tampering detection and localization compared to encryption and digital signatures [4].

Digital watermarking has the potential to provide confidentiality, origin and ownership authentication, and data integrity for exchanged medical images. Watermarking can embed patient data as robust watermarks, ensuring confidentiality and authenticity. Additionally,

watermarking enables the verification of image integrity using fragile or cryptographic hash watermarks [14].

Digital image watermarking is implemented in spatial (direct pixel modification) or frequency domains via transforms (modifying the frequency coefficients) [4].

medical image watermarking include classical, region-based (Region of Interest (ROI) and Region of Non-Interest (RONI)), and reversible methods [12, 13]. Classical techniques introduce irreversible distortions unsuitable for diagnosis. ROI/RONI spares critical regions but risks extraction if RONI is altered and limits capacity by RONI size. Reversible methods, however, fully restore originals post-extraction [13, 15].

Among reversible techniques are difference expansion (DE) [16], histogram shifting [17], and prediction error expansion (PEE) [18]. Thody and Rodriguez in [17] introduced the PEE method as an improvement to the difference expansion approach developed by Tian [16]. In contrast to DE, which uses the difference between neighboring pixels for embedding, PEE uses the prediction error [19]. In the reversible techniques, PEE excels by its ability to utilize the spatial redundancy inherent in natural images [12]. that enables high-capacity embedding with minimal distortion, making it the preferred choice among reversible algorithms [18].

Transformation domain methods such as discrete cosine transform (DCT), discrete wavelet transform (DWT), integer wavelet transform (IWT), and lifting-based wavelet transform (LWT) enhance robustness. Among these methods, LWT uniquely preserves integers, offering a reversible framework[20].

This research proposes a novel reversible watermarking scheme that integrates LWT with adaptive PEE to ensure medical image integrity, authenticity, patient confidentiality, and tamper detection/localization/recovery capabilities.

1.2 Problem Statement

Medical images and electronic patient records (EPR) are critical data that require protection from unauthorized access, tampering, and misuse. Digital watermarking is used to ensure the security of these data. However, current watermarking schemes suffer from poor robustness, limited reversibility in recovering the original image, and inadequate ability to accurately detect, localize, and recover tampered regions, and error-prone ROI selection.

1.3 Research Questions

This research aims to answer the following questions:

1. To what extent can the integration of adaptive PEE with (LIWT) improve the robustness of reversible watermarks against common image processing attacks compared to existing techniques?
2. What mechanisms can effectively detect, locate, and recover tampered regions in medical DICOM images?
3. What is the optimal trade-off between embedding capacity (bpp) and imperceptibility (measured by PSNR and SSIM) achieved by the proposed region-adaptive embedding strategy (1-bit in ROI, 2-bit in RONI) for brain DICOM images?
4. To what extent does the proposed scheme preserve diagnostic quality while ensuring confidentiality, authenticity, and integrity of medical data in compliance with HIPAA and DICOM standards?

1.4 Objectives of the Research

The main objectives of this study are:

1. To design and evaluate a novel reversible watermarking scheme for DICOM medical images that enhances robustness against attacks, achieves full reversibility, and improves tamper detection accuracy.
2. To develop and validate a region-adaptive embedding strategy that preserves diagnostic quality by embedding 1 bit per coefficient in the Region of Interest (ROI) and 2 bits per coefficient in the Region of Non-Interest (RONI).
3. To implement and assess a tamper localization and recovery mechanism for the ROI using block-level SHA-256 hashing, enabling accurate detection and restoration of tampered regions.

4. To comprehensively evaluate the proposed scheme against existing techniques in terms of embedding capacity (bpp), imperceptibility (PSNR, SSIM), reversibility, and robustness against various attacks (Gaussian noise, cropping, shearing, collage).

1.5 Importance of the Research

The importance of this research lies in preserving the security of digital medical image exchange and telemedicine, contributing to protecting patient privacy, and ensuring the reliability and integrity of medical data used in diagnosis.

Chapter 2

Literature Review

2.1 Introduction

This chapter presents the theoretical bases for the proposed LIWT+adaptive PEE watermarking scheme through background review and literature analysis on reversible watermarking. which covers basic components underpinning the proposed methodology: prediction error expansion (PEE), integer wavelet transform (IWT), k-means brain image segmentation, and SHA-256 integrity verification. addition, presents a review of 16 medical image watermarking schemes, including techniques for DICOM images in telemedicine applications. The presented studies include spatial-domain fragile approaches, robust frequency-domain transforms (DWT, LWT, DCT), and hybrid ROI/RONI segmentation techniques (manual, k-means automated, SVM, and ANN-based), including basic functionalities such as EPR embedding, integrity verification, tamper detection/localization, and recovery mechanisms. Among these, 12 representative schemes [31, 53, 55-61, 63, 64, 66] are selected for detailed comparative analysis in Chapter 5, to identify the research gaps that the proposed LIWT+adaptive PEE methodology covers.

2.2 Background Review

2.2.1 Reversible Watermarking in Medical Images

Data hiding methods in digital images are classified into two main fields: steganography and watermarking methods [21, 22]. Steganography involves protecting some additional information contained in a digital cover object—the existence of which only the sender and receiver know—as confidential information is securely transmitted through an open communication channel. While digital watermarking methods are used to protect the digital cover objects themselves, identify the author of a digital object, or confirm the authenticity and integrity of this object. Digital watermarks are often used to protect the authorship of multimedia files, to control the integrity of data, and to authenticate the sources of this data.

Reversible embedding has been the focus of attention of researchers in the past few years. Reversible embedding is a data hiding technique where the embedded data can be extracted and the original cover object (image, audio, or text) can be restored to its original form without any loss or degradation. This particular feature is increasingly important for

those applications that need some kind of intelligent processing of images that contain embedded data, such as classification and pattern recognition, where the many advanced image processing algorithms, especially those based on machine learning, are highly sensitive to even minor changes in the input image data, where Introducing distortions through irreversible data embedding can significantly degrade the performance of these algorithms; the intelligent processing algorithms may make incorrect decisions based on the altered image, and thus the only solution is to extract the embedded data and restore the cover image to its original form. Reversible watermarking enables complete recovery of medical images, allowing embedding of patient data for authentication while preserving diagnostic quality; its flexibility enables further processing after watermark removal, making it essential for ensuring the integrity and authenticity of sensitive medical data. [23-26] .

The goals of reversible watermarks for medical images include:

- Confidentiality: Ensuring the embedded patient information remains confidential and accessible only to authorized parties.
- Integrity: Maintaining the integrity of the watermarked medical images, enabling detection of any modifications, and verifying they have not been altered.
- Authentication: Providing a means to authenticate the origin and integrity of the medical images.
- Tamper Detection, Localization, and Recovery: provides a robust mechanism to detect and localize unauthorized modifications to digital images, protecting their integrity from tampering and recovering them. [24]

Reversible data embedding works with both spatial domain and frequency coefficients [27, 28]. There are many reversible watermarking techniques; however, error expansion is the most widely used method for reversible watermarking. The general idea is to calculate the error, which can be obtained directly from the difference between pixels or samples, or it can be the error between a sample and its expected value for example [29]. error expansion methods in the spatial domain that can be classified into Difference Expansion (DE), Interpolation Error Expansion, and Prediction Error Expansion (PEE). While frequency

domain approaches offer distinct advantages for securing medical images. The use of techniques like the Discrete Fourier Transform (DFT) or Integer Wavelet Transform (IWT) provides increased robustness against attacks and distortions compared to spatial domain methods. Moreover, frequency domain watermarking allows for imperceptible embedding by targeting less sensitive frequency bands, preserving the diagnostic quality of medical images [5][6][7]. frequency domain approaches perform an integer wavelet transform (IWT) on the image, they then embed the watermark in the high-frequency wavelet coefficients using error expansion as in [30] and [33].

2.2.2 Prediction Error Expansion

Prediction Error Expansion (PEE) is a reversible watermarking technique that has gained significant attention due to its efficiency and wide applicability. This method exploits the correlation among adjacent pixels to embed information bits into the prediction errors. PEE enables reversible data hiding with low distortion and high embedding capacity, without requiring a pixel site map—which makes it more efficient and practical for real-world applications. PEE is also distinguished by its ability to estimate the minimum distortion required for embedding, ensuring that the embedded data does not introduce noticeable distortions to the host image. PEE has been shown to be highly effective in various image watermarking applications, including medical imaging and digital forensics. Its ability to embed data reversibly makes it ideal for applications where data integrity and authenticity are paramount. The use of PEE, along with transform domain methods, enhances the security, robustness, and imperceptibility of watermarks in medical imaging, ensuring the integrity and authenticity of sensitive medical data.[16, 19, 30-32].

2.2.3 Integer Wavelet Transform (IWT)

IWT is used in reversible watermarking methodologies to protect medical images. IWT works as a mathematical tool to analyze images at multiple resolutions by decomposing them into different frequency sub-bands as shown in figure 1. Unlike traditional wavelet transforms that rely on floating-point arithmetic, IWT works with integers, making it well-suited for lossless applications. This feature ensures that the embedding and extraction

operations preserve the integrity of medical images by avoiding floating-point rounding errors [19].

Here is a summary of the Integer Wavelet Transform (IWT) algorithm for 2D images using the lifting scheme in basic step [33]:

Step 1. The image matrix is separated odd and even columns. Frequency sub-bands, HF (high-frequency components), and LF (low-frequency components) are calculated using the following equations:

$$HF = odd(\mathrm{i}, j) - even(\mathrm{i}, j), \tag{2.1}$$

$$Lf = \mathrm{even}(\mathrm{i}, j) + \frac{Hf}{2} \tag{2.2}$$

Step 2. LF_{even} and LF_{odd} show the even and odd rows of LF. HF_{even} and HF_{odd} are also even and odd rows of HF. First level decomposition of the image is calculated as follows:

$$HH = HE_{odd} - HF_{even'}$$

$$HL = HF_{even} + \frac{HH}{2} \tag{2.3}$$

$$LH = LF_{odd} - LF_{even.}$$

$$LL = LF_{even} + \frac{LH}{2}$$

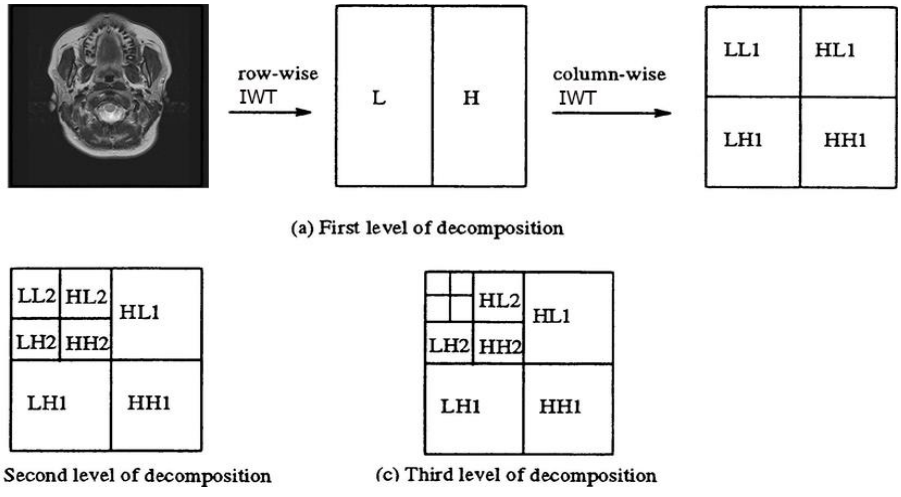


Figure 1 decomposition of Two-dimensional image using LIWT

2.2.4 Brain Image Segmentation Using K-means

Clustering methods—unsupervised machine-learning segmentation methods—divide an image into groups of pixels with similar densities without the need for training images. They use available image data to train themselves; Segmentation and training are performed concurrently by iterating between data clustering and estimating the properties of each tissue[34]. K-means clustering is the most common clustering method due to its ease of understanding and implementation. It is used in medical image segmentation to separate regions of interest from the background, particularly in brain image segmentation, where the image is divided into three main tissues: white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF) [35, 36]. This method is preferred for its simplicity, computational efficiency, and effectiveness. However, segmentation accuracy depends heavily on selecting an appropriate number of clusters (K). The K-means algorithm features ROI extraction, which is critical in applications such as watermarking to increase embedding efficiency without affecting diagnostically important regions[37].

The equations for the K-means clustering algorithm found in the provided document can be summarized as follows:

Given a set of data points $x_1, x_2, \dots, x_n$, the goal is to partition them into k clusters, each with a centroid C_a (where a is the cluster index).

The objective function Q that K-means minimizes is the sum of squared Euclidean distances between data points and their respective cluster centroids:

$$Q = \sum_{a=1}^{k} \sum_{b=1}^{n} \|x_b^{(a)} - C_a\|^2 \qquad (2.4)$$

Where:

- Q *is the total within-cluster sum of squares.*
- K is the number of clusters.
- n is the number of points in cluster a .
- X_b^a is the b^{th} point in cluster a.
- C_a is the centroid of cluster a.

The algorithm iteratively updates clusters by assigning each point to the nearest centroid through iterative steps until stability, meaning that each point is assigned to the cluster with the closest center.

2.2.5 SHA-256 for Medical Image Integrity Verification

Hash is the output of a mathematical function that takes inputs of varying lengths and generates an output called the hash value or message summary. Hash functions are one-way operations; they produce a fixed-size, irreversible representation of the input data as output. Inferring the input data from the hash value is a challenge. If the same data is input, the hash function will produce the same hash value, regardless of the input size, as shown in Figure 2. It is computationally impossible to produce the same hash value from two different inputs or to find input data that produces a specific hash value[38]. The National Institute of Standards and Technology (NIST) designed the Secure Hash Algorithm (SHA) based on the MD4 principle, and it was published as a Federal Information Processing Standard (FIPS 180) in 1993. In 1995, a revised version called SHA-1 was released, which produces a 160-bit hash value. In 2002, the National Institute of Standards and Technology (NIST) released a revised version of the standard known as FIPS 180-2, defining three new versions of SHA with lengths of 256, 384, and 512 bits, known as SHA-256, SHA-384, and SHA-512, respectively [39]. Attacks on SHA-0 and SHA-1 have been reported, while no practical attacks on SHA-2 have been recorded to date, thanks to its improved structure and longer output lengths. SHA-256 offers a strong balance between robustness and performance [40]. SHA-256 processes data by splitting it into 512-bit blocks, each of which undergoes 64 rounds of complex calculations. SHA-256 is used to verify the integrity of medical images and detect any modifications or tampering. Any slight change to the image alters the resulting hash value; this allows for the detection of any unauthorized modifications. These hash values are used as watermarks to verify image authenticity, confirm ownership, and enhance trust and security in healthcare settings.

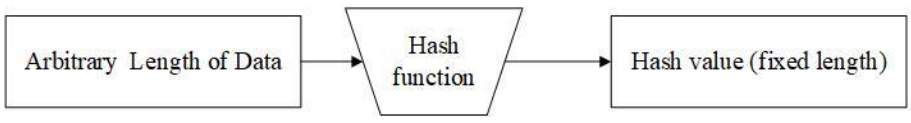


Figure 2 Hash Function

2.3 Related Literature Review

The field of reversible watermarking for medical images has rapid development to protect the diagnostic integrity of images and the confidentiality of patient data in healthcare. Despite the advanced development, there are significant research gaps that still exist that limit the reliability of these technologies. In this review, the researcher presents a critical analysis of related studies, organized according to the objectives of the proposed research. In each section, related work is analyzed, and its methodological and application limitations are identified, highlighting the specific research gaps that this work addresses.

2.3.1 Reversible Watermarking Via Wavelet Transforms and PEE

Mamatha and Ahmad [41] demonstrated the effectiveness of 5/3 and 9/7 integer wavelet transforms for reversible data hiding using the standard prediction error extension (PEE). Their Comparative analysis showed that these transforms outperform HAR waves in achieving a balance between capacity and quality, as later confirmed by a study [42]. However, this study relied in general on a static, non-adaptive PEE version and was limited to providing a comparative basis for evaluation, ignoring necessary functions in healthcare applications, such as tamper detection and robustness evaluation.

Yi et al. [43] introduce the adaptive block-level prediction-error expansion (ABPEE) method for reversible data hiding in encrypted images. The technique embeds secret data into 2x2 image blocks using pixel redundancy. It outperforms other methods in terms of overflow map size reduction and visual quality improvement. However, further research is needed to ensure robustness against potential attacks.

The authors presented in[44] a novel reversible watermarking scheme for MRI and CT brain images based on adaptive PEE. The scheme automatically divided the medical image into ROI and RONI. Patient data and the hash value of the ROI were embedded within the ROI, while data for detecting and recovering tamper were stored in the RONI. The scheme used an adaptive embedding, where it allocated one bit for each embeddable pixel in the ROI and two bits in the RONI. the scheme demonstrated the ability to identify and recover altered regions and achieved a hiding capacity of 1.5 bpp. However, the method faces some challenges. First, their approach hides data in the spatial domain, which limits the system's robustness compared to frequency-

domain techniques. The testing of robustness was limited to tampering attacks specifically cropping attacks to test its ability to detect modifications. The scheme was not evaluated against image processing attacks. Second, for verifying data integrity, the scheme employs the weak and no longer reliable MD5 algorithm; this raises security challenges. Finally, dependence on average block values does not provide enough sensitivity for detecting modifications; this is a limitation in ensuring block-level integrity.

Kapadia and Nithyanandam [45] present a secure and high-capacity reversible data hiding technique using the Integer Wavelet Transform (IWT) and Arnold Transform. The scheme divides the original cover image into four sub-bands using IWT and applies the Arnold transform to enhance visual attack resistance. A 2-bit pseudo-LSB random embedding technique is used to embed binary streams into these sub-bands. Overall, the paper introduces a reversible watermarking approach but, limited capacity due to storing location maps.

Research Gap: To the best of our knowledge, no existing study has integrated adaptive PEE with LIWT while simultaneously addressing robustness, reversibility, and ROI-based embedding for full DICOM images.

2.3.2 ROI/RONI Segmentation and Region-Adaptive Embedding

Pal et al. [20] proposed a medical image watermarking scheme primarily aimed at security and authentication, in which Support Vector Machines (SVM) are used to automatically segment breast images into small lesional Regions of Interest (ROI) and surrounding NROI regions, while the actual watermark is embedded only in the NROI using the Lifting-based Wavelet Transform (LWT). The method achieves high visual quality (PSNR around 66–67 dB and SSIM close to 1) and relatively high payload within the NROI, using a 165×165 grayscale logo embedded in 256×256 MR images. However, the scheme presents several limitations. First, the ROI is restricted to local lesional areas and does not cover the extended diagnostic region typically required by clinicians, and no watermarking, integrity verification, or tamper recovery is provided for the diagnostic ROI itself, which constrains its clinical applicability. Second, the approach does not use any cryptographic hash or tamper map, relies only on the presence of the watermark for authentication, and reports BER values only under zero-attack and dataset variations settings, without BER evaluation under specific image processing or

tampering attacks, which limits the robustness analysis in medical watermarking. Finally, although the authors describe the scheme as reversible, the reported quality metrics (finite PSNR and SSIM < 1) indicate near-lossless rather than strictly reversible behavior in the sense of perfect reconstruction required for medical RDH applications.

Bhalerao et al in [46]. proposed a reversible medical image watermarking scheme for ROI tamper detection and recovery. The scheme divides the medical image into ROI and RONI regions. A low payload of only 160 bits (0.016 bpp) is embedded in the ROI using prediction-error expansion for authentication. While the RONI carries a payload that has a size of 45,360 bits for a 512x512 image. This RONI payload is composed of the compressed ROI copy, the RONI authentication hash, overflow location maps, and ROI overflow data. The scheme achieves high imperceptibility with an average PSNR of 55 dB and SSIM of 0.99. However, it has a challenge in an integrity validation due to its use of an insecure SHA-1 hashing function. It also suffers from limitations and unreliability in processing the entire diagnostic region as a single unit, where it manually selects a small, fixed-size (100 × 100 pixel) region of interest and lacks security analysis against common manipulation attacks.

A study in [47] presented a reversible medical image watermarking method to detect and recover tampering at the ROI level. The scheme secures medical images through three main functions: storing patient data, detecting tampering, and recovering tampered regions. The design divides the image into manually defined regions using polygonal coordinates: ROI, where patient records and an ROI hash are embedded using a multi-layer quad-based difference expansion method with a capacity of up to 6 bits per pixel; the Region of Non-Interest (RONI), which uses a two-dimensional difference expansion (2D-DE) scheme to store recovery and auxiliary data; and a border region that embeds ROI and RONI coordinates and maps using least significant bits (LSB). The system achieves high visual quality for watermarked images, with PSNR values reported up to about 50.86 dB, and demonstrates effective ROI recovery with quality gains of up to +27.61 dB compared with the tampered state. However, the method still suffers from several limitations, including the use of the cryptographically weak MD5 hashing function, an insensitive tamper detection mechanism based on block averages, protection of only a limited fraction (4–12%) of the actual diagnostic region, and significant capacity overhead devoted to location maps. Moreover, the use of LSB embedding in the border

region introduces permanent and irreversible modifications to pixels that may lie within clinically relevant portions of the image, there is no clarification of how fractional coefficients from the transform domain are handled, and there is no robustness evaluation against unintentional or common image manipulation attacks such as Gaussian noise.

Research Gap: Current methods use manual, fixed, or incompletely segmented ROI, lack automated segmentation, and lack adaptive embedding according to region.

2.3.3 Confidentiality, Authenticity and Integrity of Medical Images

Shi et al. [48] proposed a region-based reversible watermarking method for medical images, aimed at privacy protection and integrity checks. It creates three watermark types: authentication, restoration, and privacy and showed good efficiency on large grayscale images (960×960) with high capacity and PSNR, though capacities were theoretical. Still, it has key flaws: reconstruction uses neighborhood averages, which blur high-contrast areas and fail to restore details accurately (no quantitative metrics provided, and tests avoided edges). It skips DICOM/PACS compatibility, metadata handling, and diagnostic validation. The location map cuts effective capacity, and there's no testing against frequency/amplitude attacks or tamper localization typical weaknesses of spatial-domain methods.

Hurrah et al. [49] present a scheme for medical image authentication and tamper detection in e-health environments. The scheme is based on dividing grayscale images into 4x4 blocks, where patient data (with a 16 KB payload) is embedded by modifying the average pixel value in each block. To ensure recovery and tamper detection, the modulation factor is derived during the embedding process and stored in the least significant bits (LSB) of a reserved pixel within each block, allowing the modification effect to be reversed for the remaining pixels. This method achieved more than 47 dB for PSNR. Despite enhanced security through multiple encryption techniques (Arnold mapping, grayscale, and AES-128 encryption), the scheme's reliance on spatial scale makes it vulnerable to signal manipulation attacks. Furthermore, the destructive modification used to extract the modulation factor from the LSB results in distortion of the reserved pixel, rendering the system semi-reversible rather than fully reversible. This compromise on the accurate recovery of some pixels is necessary to avoid the high cost of a location map that consumes available payload.

Geetha and Geetha [50] discuss the use of reversible data hiding (RDH) to improve medical data security and electronic patient record (EPR) privacy in medical images. They employ the rhombus mean interpolation technique to embed higher data capacity while maintaining image quality. The embedded data includes a checksum for tamper detection and content authentication. The method is evaluated using metrics like peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). this study does not discuss how the proposed approach handles the Region of Interest (ROI) within the medical images, need for segmentation and frequency domain robustness.

Li et al. [51] propose a waveform-based reverse watermarking algorithm for authenticating medical image content. The image is divided into 2 × 2 blocks, and waveforms are performed on each block. Two watermarks are embedded: one for authentication, embedded in the smallest bit of the approximated array of parameters, and the other for the properties, embedded in the second and third bits of the same parameters. Chaotic mapping is used to enhance security and prevent random embedding. The algorithm employs various levels of inspection to detect tampering and exhibits high quality after embedding and approximate recovery. Although the frequency domain is used, the scheme does not provide a comprehensive assessment of the robustness of the watermark against common image processing attacks. Additionally, the scheme suffers from high computational complexity due to the operations required to process the fractional part of the real values produced by the discrete wavelet transformation used.

Nguyen and V [52] proposed a reversible image authentication scheme for digital images. The scheme divides the host image into 3x3 non-overlapping blocks and generates an authentication code for each block. The authentication code is adaptively embedded using prediction error expansion. The scheme lacks details on computational complexity, security considerations, and evaluation methodology, suggesting further analysis. while the scheme aims to provide reversible authentication, it faces challenges in ROI handling, manual segmentation, limited capacity due to storing location maps, lack of frequency domain robustness evaluation, and unquantified tamper localization accuracy, which constrain its overall performance and applicability.

Research Gap: To the best of our knowledge, no integrated schemes for DICOM images and their metadata, and ensures EPR confidentiality, image authenticity, and ROI integrity with tamper localization and recovery.

2.3.4 Performance Evaluation: Capacity, Quality and Robustness

Agrawal [53] presents a reversible data hiding technique for medical images using the integer-to-integer wavelet transform and histogram-bin-shifting. The technique divides images into blocks and evaluates their smoothness using entropy. The watermark is embedded in the detail sub-bands using the wavelet transform. Experimental results demonstrate higher PSNR values compared to existing methods, making it suitable for securing medical images. However, the article lack of discussion on tamper detection capabilities of the proposed IWT-domain watermarking technique, and it provides limited information on how the technique handles and preserves the Region of Interest (ROI) in medical images.

In [54], a spatial domain scheme for hiding electronic patient information (EPI) in medical images was presented. The method is based on Enhanced Neighbor Mean Interpolation (ENMI), where each 2x2 block from the original image is expanded into a 3x3 block. Within each expanded block, five interpolated pixels are used to embed a total of 4 bits of secret EPI data and 1 bit of an authentication code, achieving a fixed embedding capacity of 327,680 bits (0.55 bpp). The system demonstrated high fidelity, with an average PSNR of 46.94 dB and SSIM of 0.984 for the stego images, making visual detection of changes difficult. However, the scheme has notable limitations: it lacks a mechanism to recover tampered areas, only detecting them, and shows poor robustness under common image processing attacks, especially Additive White Gaussian Noise (AWGN), which caused a Bit Error Rate (BER) as high as 40.64%. Furthermore, its fixed embedding capacity does not adapt to local image characteristics.

The authors in [55] presented an advanced algorithm for embedding a dual watermark in DICOM images for copyright protection. The scheme was tested on 512 × 512 ultrasound images of the liver from three different patients. The algorithm employs multiple analysis techniques, starting with a lifting-based wavelet transform (LWT) to decompose the host image into three levels of frequency bands. It then uses Schur analysis, discrete cosine transform (DCT), and single-value decomposition (SVD) to process the selected frequency bands. Two

watermark images, each containing the corporate logo and patient data, were embedded at 256 × 256 pixels, with scaling factor optimization using the Firefly algorithm. The system achieved high PSNRs exceeding 40 dB on average, demonstrating a good balance between robustness and invisibility. However, the algorithm suffers from several limitations, the most important of which are the limited sample size (only 3 images), the lack of complete reversibility due to information loss during multiple decomposition processes, and its inability to detect local manipulation or recover original data, in addition to the high computational complexity resulting from the complex integration of multiple techniques (three-level LWT with DCT, SVD, and Schur), which limits its suitability for real-time applications. In addition, there is a lack of encryption mechanisms for sensitive medical data before embedding.

Liu et al. [56] confirm medical image integrity and authenticity through a hybrid SLT-SVD transformation. The technique divides regions of interest (ROI) and regions of non-interest (RONI) into segments and adds a watermark throughout the entire image to guarantee its legitimacy. The technique combines the integer wavelet transform (IWT) with the bit-plane technique (BTC) to recover information about manipulation. The scheme relies on manual division of the image into ROI and RONI for watermark generation, an automated and robust segmentation method would enhance its practicality and scalability.

In [57], Bamal et al. presented a hybrid system for reversible watermarking with tamper detection and recovery for medical images. Artificial neural networks (ANNs) were used to divide the image into ROI and RONI; hybrid watermarking uses SLANLT transform (SLT) for frequency-domain embedding and RS vector modulation for spatial-domain embedding in 512×512 8-bit images. SHA-3, AES, and LZW are used for reliability and confidentiality. The total average embedding capacity was 62,119 bits, of which 19,361 bits were for ROI and 42,925 bits for ROI. It achieved an average PSNR of 56.28 dB and resistance to various attacks, such as Gaussian noise (AGN=0.0008). However, the scheme suffers from limitations in tamper detection and recovery with the expanded diagnostic regions relied upon by clinicians for comprehensive diagnosis, as the evaluation was restricted to very small ROI. Furthermore, the study failed to clarify the capacity of side information required to achieve full reversibility, which consumes available capacity. Moreover, the proposed method included a preprocessing

step to enhance contrast, which contradicts the principle of full reversibility and alters the original image values.

Research gap: To the best of our knowledge, most studies lack comprehensive evaluation across robustness, capacity, quality, and reversibility metrics. Moreover, a critical gap persists between simplified methodological steps (relying on small, manually selected regions) and clinical applicability, which requires processing full DICOM data with metadata while preserving the integrity of the entire diagnostic area as a single unit.

2.4 Summary

This chapter covered the theoretical basis, which reveals several limitations in current reversible watermarking schemes: spatial methods sacrifice recovery for capacity, frequency approaches lose reversibility for robustness, and ROI techniques have challenges with automation and scalability. where the background confirms PEE+IWT+SHA-256+k-means are the most effective components, while the literature review of 16 schemes identifies critical gaps in EPR/DICOM integration directly and confirms the proposed LIWT+adaptive PEE solution, evaluated against 12 key methods in Chapter 5

Chapter 3

Methodology

3.1 Introduction

This chapter provides an overview of the systematic steps involved in designing a new reversible watermarking scheme for the transformation domain. It is supported by DICOM medical image files. The objectives of this scheme are to achieve a balance between diagnostic accuracy, high embedding capacity, and data security. The approach is multi-stage, combining a lifting-based integer wavelet transform (LIWT) with adaptive prediction error expansion (PEE) technology to embed reversible watermarks. Intelligent segmentation defines a region of interest (ROI) over non-diagnostic regions (RONI), improving control over watermark embedding. The framework includes protocols for tamper detection, integrity verification, and precise localization of alterations. considered with HIPAA and DICOM requirements, image transparency, lossless restoration of original images, and high accuracy in tamper localization are considered.

3.2 Methodology Overview

The following two figures illustrate the embedding and extraction process of the proposed scheme:

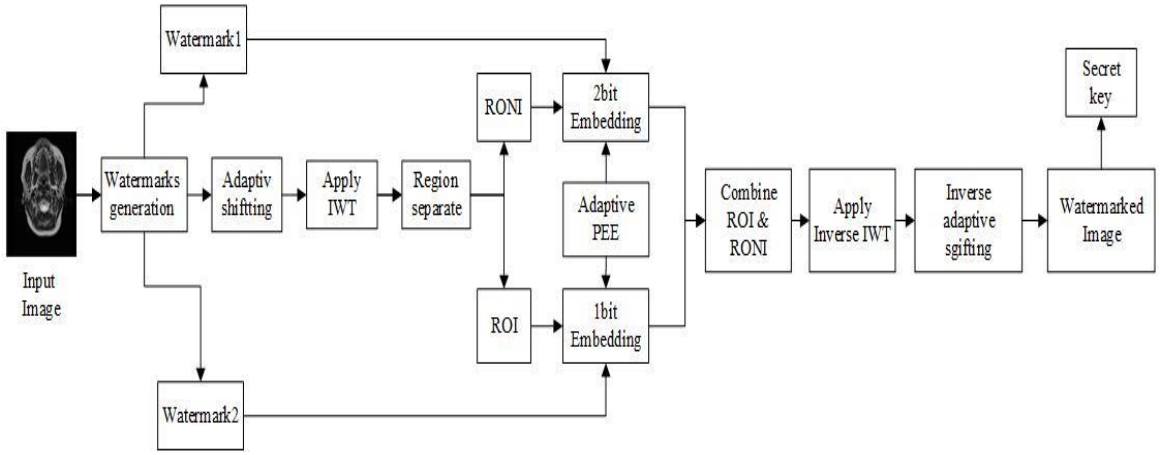


Figure 3 Embedding process

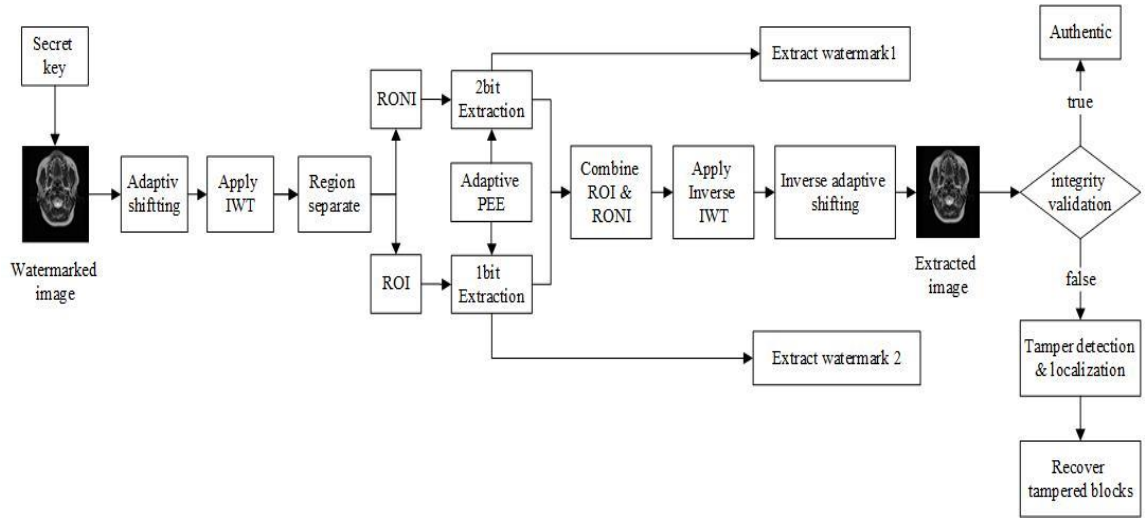


Figure 4 Extraction process

3.2.1 Reading DICOM Images and Extracting Metadata:

We begin by loading medical images $C(\mathrm{i}, j)$ in DICOM format and extracting all associated information, this step is crucial for ensuring accurate data documentation and forms the foundation for subsequent analyses.

3.2.2 Preprocessing and Analysis

This phase involves a systematic series of steps to prepare and analyze medical images with precision, as follows:

- **Image Segmentation Using K-means Algorithm**:

Based on the theoretical foundation established in Section 2.2.4, we employ the K-means clustering algorithm with K=4 to automatically segment the image into homogeneous regions. K-means was selected for this task due to its established effectiveness in segmenting brain MRI and CT images into three main tissue classes (white matter, gray matter, and cerebrospinal fluid) plus the background, as well as its computational efficiency, simplicity, and proven performance in medical image segmentation. These characteristics make it particularly suitable for the automated ROI extraction required in this watermarking framework, where it is expected to reliably separate the diagnostically relevant head region (ROI) from the non-diagnostic background (RONI). The detailed implementation, including parameter selection and morphological post-processing, is presented in Section 4.2.2. While K-means has known limitations—such as sensitivity to initialization and the assumption of spherical clusters—its adequacy for the specific task of isolating the head region is demonstrated in the experimental results (Section 5.2.1, Figures 8 and 9), where it successfully segmented ROI areas comprising 26-50% of the image across all tested modalities.

- **Preventing Underflow in DICOM Medical Images**

DICOM images are typically stored with pixel values of 12- or 16-bits depth [58], providing a wide numerical range. However, the actual pixel values are often much lower than the maximum possible value, making overflow issues unlikely. Due to the proximity of some pixel values to zero, underflow remains possible. Therefore, as part of the research methodology and to avoid the underflow phenomenon, an image histogram adaptive technique was proposed. In this approach, suitable offsets are calculated separately for both the ROI and RONI regions according to the characteristics of both the image and the watermark.

3.2.3 Multi-Component Watermark Generation

The process of generating multi-component watermarks in medical images consists of a diagnostic component and a security component. These are integrated to validate diagnostic integrity, enhance security, and maintain data confidentiality.

in our approach, the two watermarks are generated by converting the extracted DICOM file attributes into a binary string, which is then combined with the ROI hash value to form the first watermark. This watermark is embedded into the ROI coefficients of the medical image. The second watermark is created by merging the RONI (Region of Non-Interest) hash with both the DICOM-compressed ROI data and the tamper-detection data derived from ROI block partitioning. Prior to embedding, both watermarks are encrypted using the AES-128 algorithm. AES-128 was selected due to its proven security, widespread adoption in medical applications, and optimal balance between robustness and computational efficiency, enhancing security and preventing unauthorized access. Recent research by Lata et al.[1] confirms that AES-128 achieves encryption rates up to 21.8 Mbps while maintaining high security metrics (PSNR > 50 dB in CBC mode, NPCR > 99.5%) for medical images, making it suitable for real-time healthcare applications. The encryption keys are managed separately as part of the key generation process (Section 4.2.12).

3.2.4 Integer Wavelet Transform (IWT)

In this proposed scheme, to increase robustness, the integer wavelet transformation based on the lifting scheme was selected as the fundamental processing step for the proposed reversible watermarking approach. Specifically, The CDF 5/3 integer wavelet transform was implemented using the identifier 'cdf2.2' in a MATLAB environment. To verify technical accuracy, the resulting lifting scheme was analyzed, confirming that its parameters [-0.5, -0.5] and [0.25, 0.25] perfectly matched the standard lifting algorithm for the 5/3 transformation[59]. This also conforms to the standard specification for the CDF 5/3 transformation used in the JPEG 2000 lossless compression standard [60]. This confirms that the representation the 'cdf2.2' software corresponds to the theoretical 5/3 transformation that the previous study [41] indicated as one of the most efficient transformations for reversible data hiding.

apply the IWT to the medical image using a lifting scheme, decomposing the image into four sub-frequency bands: approximation (CA), horizontal (CH), vertical (CV), and diagonal (CD).

3.2.5 Adaptive Capacity Thresholds

The expected values of the frequency coefficients are computed, followed by the calculation of prediction errors for these coefficients. This analytical process determines the available embedding capacity. Further details of each step will be elaborated in subsequent sections.

3.2.6 Data Embedding Process

The proposed approach implements a reversible watermarking system based on the integration of Adaptive (PEE) and (LIWT). In this method, the cover image in DICOM format C(i,j) with dimensions M×N is processed using LIWT, generating four sub-bands (CV, CD, CH, CA), $C \overset{\text{IWT}}{\rightarrow} \{CA, CH, CV, CD\}$, each of size M/2×N/2. Within this framework, an adaptive PEE algorithm was employed with varying embedding strengths: one bit is embedded in the Region of Interest (ROI), while two bits are embedded in Regions of Non-Interest (RONI) of the wavelet transform coefficients. this adaptive embedding strategy maintains diagnostic accuracy in critical regions (ROI) while maximizing payload capacity in non-critical regions (RONI), achieving an optimal balance between data-hiding requirements and reconstructed image quality.

3.2.7 Key Generation

Following completion of the embedding process of the embedding process, a secret key is generated to enhance system security, ensuring that only authenticated entities can extract the watermarks and access protected patient information.

3.2.8 Data Extraction Process

For extraction and recovery, the system requires two primary inputs: the watermarked image and the original secret key. The LIWT decomposition is reapplied to get the frequency sub-bands, followed by ROI/RONI localization through binary masking. The watermark extraction process retrieves embedded bits from each sub-band while maintaining coefficient integrity, with special attention to the adaptive PEE scheme's dual-strength embedding.

Extracted watermarks undergo AES decryption when applicable. Image reconstruction employs Inverse LIWT (ILWT) to perfectly reassemble the sub-bands, with subsequent pixel-value correction to reverse any modifications introduced during the embedding process.

3.2.9 Summary

This chapter provided an overview of the methodology used for a reversible watermarking system for DICOM images. This system is based on the Lifting-based Integer Wavelet Transform (LIWT) and Adaptive Prediction Error Expansion (PEE) to embed watermarks while preserving diagnostic image quality. The system divides images into diagnostic regions (ROI) and non-diagnostic regions (RONI) to control watermark embedding. Two cryptographic watermarks are embedded to ensure data integrity and confidentiality, while image authenticity is maintained through tamper detection and localization mechanisms. This methodology ensures full restoration of the original image and complies with healthcare security standards.

Chapter 4

Proposed Methodology: Design and Implementation

4.1 Introduction

This chapter provides a robust mathematical and methodological framework for a reversible watermarking technique for medical images, according to a literature review and the issues identified in previous chapters. Depending on this challenge, we present the basic mathematical models that describe the embedding and extraction processes. This study advances the field of reversible watermarking for medical images, addressing several key challenges. First, it shifts from spatial domain techniques to frequency domain solutions, guaranteeing the exact recovery of the original image through the use of the IWT. Second, it improves the balance between capacity and accuracy by integrating the IWT with PEE using adaptive thresholds, allowing precise restoration of every bit in the original image after extraction. The system also includes intelligent segmentation of the medical image to preserve diagnostic regions and enhance embedding efficiency.

The proposed system can be designed in several stages: initial image preprocessing, watermark generation, data embedding in the frequency domain, authentication, data lossless extraction and recovery of the original image, and tamper detection, localization, and recovery. These improvements make the system ideal for sensitive medical applications where lossless preservation of diagnostic data is critical.

4.2 Design and Implementation of Proposed Scheme

In this section, we explain the proposed methodology, in terms of the embedding and extraction processes of the reversible watermarking algorithm, as well as the authentication and integrity verification of medical images.

4.2.1 Preprocessing and Analysis

This stage includes a series of operations performed on the input medical images $C(\mathrm{i}, j)$, which are in DICOM format, to improve image quality and prepare the data for subsequent effective analysis, as shown below:

4.2.2 Image Segmentation Using K-means Algorithm

- The proposed methodology integrates K-means initialization with polygon simplification to achieve effective segmentation of MRI and CT scans of brain images into four well-

defined anatomical regions. Initially, the original medical image $X \in R^{hxw}$ is converted to an appropriate digital format (e.g., uint16) where h and w represent the image height and width, respectively. The segmentation employs the K-means clustering algorithm with $k = 4$, reflecting the four main brain tissue classes. The cluster centroids are iteratively updated by assigning each pixel i to the nearest centroid, thereby grouping the pixels into clusters based on their closest centroids until convergence.

- Post-processing includes filling internal holes via morphological filling and smoothing boundaries through two iterations of erosion using a diamond-shaped structuring element. The largest connected component, representing the region of interest (ROI), is selected using connected-component analysis, and its relative area is quantified as:

$$\text{ROIsize} = \frac{\text{maxArea} \times 100}{h \times w} \tag{4.1}$$

The ROI boundary C is extracted and simplified through polygon decimation algorithms:

$$C_2 = \text{DecimatePoly}(C, \epsilon),$$

reducing the number of contour points while preserving geometric fidelity. Subsequently, a binary mask BW_{ROI} is generated using the simplified polygon vertices via

$$BW_{\text{ROI}} = \text{roipoly}(X, x_i, y_i) \tag{4.2}$$

and the corresponding pixel data is isolated as:

$$X_{\text{ROI}} = X \odot BW_{\text{ROI}}, \tag{4.3}$$

Where$\odot$ denotes element-wise multiplication. The polygon's binary vector representation size and total side information size support efficient encoding and after processing. As shown in Figure 5, through the segmentation process, the full head mask is identified as a region of interest (ROI) containing diagnostic information of critical importance to the clinician, while the non-interest region (RONI) comprises the black background of the medical image, completely devoid of any vital diagnostic data. These segmentations ensure that the region of interest is treated as a single, integrated unit for medical analysis and diagnosis. This integrated methodological and mathematical

framework ensures precise segmentation and compact representation of brain regions of brain regions for medical images.

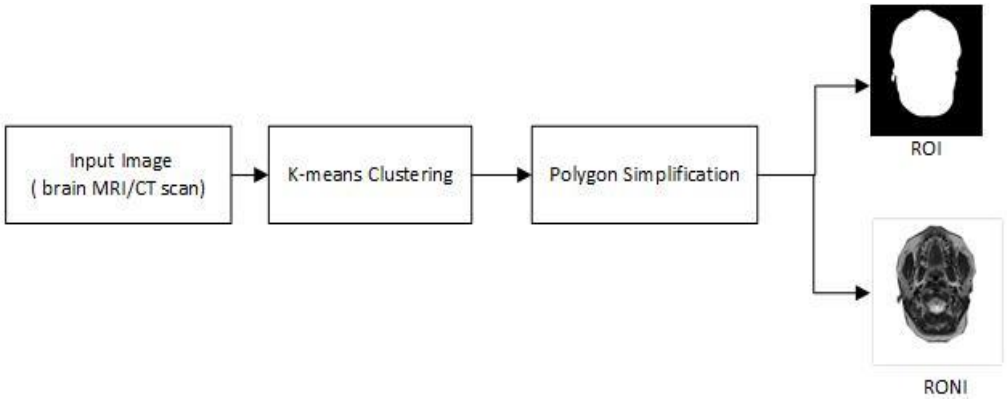


Figure 5 segmentation brain image into ROI and RONI

4.2.3 Adaptive Shifting Methodology to Prevent Underflow

Because some pixel values are close to zero, underflow may occur when embedding data into these pixels. To address this, an adaptive image shift scheme is proposed that computes appropriate shifts separately for the region of interest (ROI) and the region of no interest (RONI). This technique dynamically adjusts the shift values according to the image and watermark characteristics to prevent data underflow in both regions. The shift values are determined based on image properties, including the ratio of the ROI to the overall image size and parameters related to the watermark or image that influence the shift value. Adaptive transformations for both ROI and RONI are then calculated and applied to the original image pixels prior to embedding.

The final watermarked image I' is obtained by shifting pixels of the original image I as follows:

$$I'(i,j) = \begin{cases} I(i,j) + \Delta_{ROI}, & if mask(i,j) \neq 0 \ (pixel\ in\ ROI) \\ I(i,j) + \Delta_{RONI}, & if mask(i,j) = 0 \ (pixel\ in\ RONI) \end{cases} \quad (4.4)$$

where mask is a mask indicating the region of each pixel, Δ_{ROI} is the shift value in the ROI, and Δ_{RONI} is the shift value in the RONI.

4.2.4 Generate the Watermarks

In the proposed watermarking scheme, two distinct watermarks are generated for adaptively embedding into the medical image. The first watermark is embedded in the Region of Non-Interest (RONI), while the second watermark is embedded in the Region of Interest (ROI). This separation ensures that the critical diagnostic information within the ROI is preserved with integrity, while the RONI carries additional watermark data without affecting image quality. Figure 6 shows the process of generating a watermark.

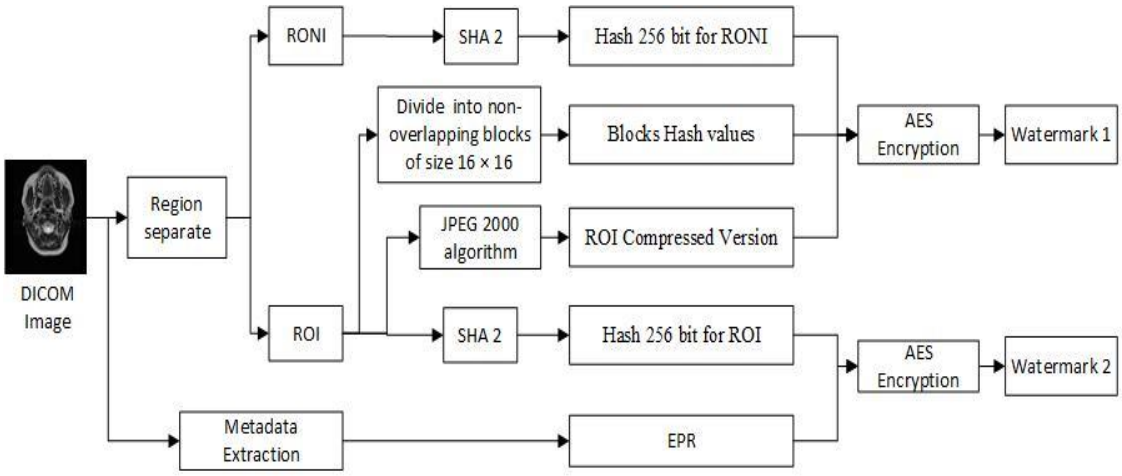


Figure 6 Watermarks generation

4.2.5 The First Watermark

The first watermark is embedded in the RONI to ensure the integrity of the diagnostic area of interest. It consists of three different data components: the first is a 256-bit RONI hash value; the second is a compressed version of the area of interest used for data recovery; and finally, the hash values of the blocks within the area of interest are included to ensure the detection of any tampering.

In order to ensure the security, this watermark is encrypted using the AES algorithm before being embedded in the RONI of the frequency coefficients.

Step 1: Compute SHA-256 Hashes for ROI and RONI

- After determining the ROI and RONI regions using equation (4.3), secure SHA-256 hash values are computed separately for each region.
- Each SHA-256 hash value is 256 bits long and is then converted into a binary vector.
- The hash corresponding to the RONI region is embedded within the RONI to verify its integrity.

Step 2: Block Division and Reduced Hash Computation for ROI

- The ROI is divided into non-overlapping blocks of size 16 × 16. This division allows for block-level authentication, tamper detection, and independent recovery.
- total blocks compute by $h * w$ where:

$$h = \left\lceil \frac{Height_{ROI}}{16} \right\rceil, \qquad w = \left\lceil \frac{Width_{ROI}}{16} \right\rceil \tag{4.5}$$

- For each block Block $_{i,j}$, compute:

$$H_{i,j} = SHA256\left(\text{Block}_{i,j}\right) \tag{4.6}$$

- Apply hash reduction for each hash $H_{i,j}\,(256\text{ bits})$: split the 256-bit hash into two halves of 128 bits each.

First half divided into 8 groups, each 16 bits, compute OR across each group:

Second half divided similarly, compute XOR.

- To obtain the final reduced hash of the block, all the reduced hashes are collected into an array representing the block hash and this process is repeated for all blocks.
- The calculate Reduced Hash process has been illustrated graphically in Figure 7.

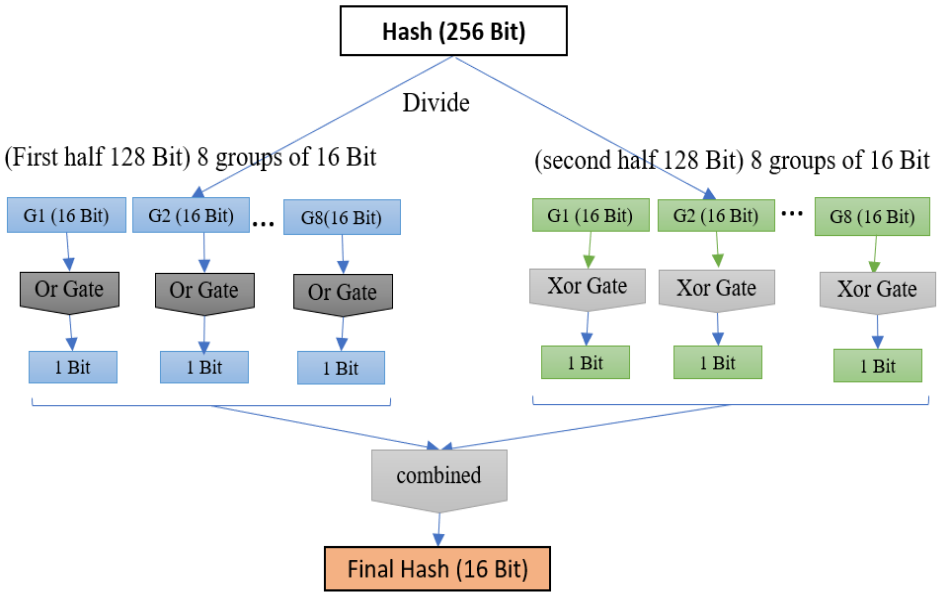


Figure 7 Hash Value Reduction Function

Step 3: Compression of ROI Data

- A compressed version of the Region of Interest (ROI) is embedded within the Region of Non-Interest (RONI). This embedded compressed data facilitates the accurate restoration of the ROI if any tampering occurs in the ROI region.
- According to the guidelines outlined in Section 8 from Part 5 of the DICOM standard [33], the pixel data of the (ROI) is compressed using the JPEG 2000 algorithm, which provides more efficient compression with less distortion, making it particularly suitable for medical imaging where preserving image integrity is critical for accurate diagnosis [61]. The compression ratio is carefully selected based on the imaging modality to strike a balance between efficient data reduction and preservation of diagnostic information. Specifically, a compression ratio of 10 is used for MRI images, while a higher ratio of 15 is applied for CT images. These values align with DICOM guidelines and optimize the trade-off between compression efficiency and the need to retain essential medical details for reliable diagnosis.

4.2.6 The Second Watermark

The second watermark is a security component embedded in the interest region. It contains two distinct data components: the first is the electronic patient record (EPR), which ensures

confidentiality and authenticity; the second is the ROI data hash signatures, a unique digital fingerprint that aids in detecting tampering or modification and verifies the integrity of the ROI data. To ensure security, this watermark is encrypted using the AES algorithm before being embedded into the ROI frequency coefficients.

Step 1: SHA-256 Hash for ROI

- The hash value calculated during the RONI embedding step is used for the ROI region and embedded within the ROI to ensure its integrity.

Step 2: Metadata Extraction:

Metadata fields are extracted from DICOM images; these are attributes that must be protected to ensure a minimum level of confidentiality, according to Section 15 of the DICOM standard[33, 62]. These attributes include patient personal information, technical specifications of the images, and details of the imaging device. Protecting these attributes is essential for maintaining data confidentiality in medical imaging systems and preventing information leakage.

- Fields extracted from DICOM metadata are defined as:

$$\mathcal{A} = \{a_1, a_2, \dots, a_n\} \tag{4.7}$$

where each a_{i} represents a specific DICOM attribute, like patient name, date of birth, gender, etc.

- Given the input DICOM metadata structure M , the extraction function f maps each attribute $a_i \in A$ to its corresponding value v_{i} if it exists:
- Each DICOM metadata attribute is verified and its value extracted.

$$v_i = f(a_i, \mathbf{M}) = \begin{cases} \mathbf{M}(a_i) & \text{if } a_i \in \text{fields}(\mathbf{M}) \\ \emptyset & \text{otherwise} \end{cases} \tag{4.8}$$

- All extracted attribute-value pairs are organized into a structured record, which is then converted into a binary data stream for subsequent processing. This data stream is encrypted using the AES algorithm to provide an additional layer of security before being embedded in the ROI of the frequency coefficients.

4.2.7 Integer Wavelet Transform (IWT)

- Apply the IWT to the medical image based on a lifting scheme, to decompose the image into four sub-frequency ranges: approximation (LL), horizontal (LH), vertical (HL), and diagonal (HH) bands.
- Define ROI and RONI: In the proposed scheme, Equation (2.1) is used to segment the approximation domain (LL sub-band) of the input medical image of the head of the patient. This segmentation divides the LL sub-band into two parts—ROI and RONI—and generates a corresponding binary mask. This mask is subsequently applied to segment the remaining sub-bands accordingly.
- The subcomponents are combined by combining the transformation results extracted from the original image into a unified data structure. The combination process is performed by first ordering the approximation coefficients, followed by the horizontal, vertical, and diagonal detail coefficients, preserving the hierarchical structure extracted from the waveform transformation.

$$\text{sub} - \text{bands}_{merged} = [CA, CH, CV, CD] \tag{4.9}$$

- Binary masks are merged in the same order as the sub-bands IWT coefficients.

$$BW_{merged} = [BW_{CA}, BW_{CA}, BW_{CA}, BW_{CA}]$$

In this way, the coefficients and -binary masks are grouped in an organized manner for easier handling during subsequent processing.

4.2.8 Determine the Available Embedding Capacity of ROI and RONI

Step1: Computing prediction error for each point

The matrix generated from the equation (14), denoted as $I(x, y)$, is scanned in raster-scan pixel-by-pixel order, excluding the first row and first column. Each coefficient is then predicted using Equation (2.10):

$$I'(x, y) = \left[\frac{I(x-1,y) + I(x,y-1)}{2}\right] \tag{4.10}$$

calculated the prediction error using (2):

$$e = |I(x,y) - I'(x,y)| \tag{4.11}$$

Step **2:** Classification of Coefficients into Expandable, Shiftable, and Non-changeable Sets

Using various adaptive thresholds T_1 for ROI and T_2 for RONI, where the threshold values are initially set to zero and gradually increase, coefficients are classified into three mutually exclusive sets for each region

- For coefficients located in RONI where $E_{RONI} = \{(i,j)|BW(i,j) = 0\}$:

$$\begin{cases} E_{RONI}, & \text{if } e(i,j) \in [-T_2, T_2] \\ S_{RONI}, & \text{otherwise} \end{cases} \tag{4.12}$$

- For coefficients located in ROI where $E_{ROI} = \{(i,j)|BW(i,j) = 1\}$:

$$\begin{matrix} E_{ROI}, & \text{if } e(i,j) \in [-T, T] \\ S_{ROI}, & \text{otherwise} \end{matrix} \tag{4.13}$$

where S_{RONI} is the set of shiftable coefficients in the Region of Non-Interest (RONI), and E_{RONI} is the set of expandable coefficients in the RONI, S_{ROI} is the set of shiftable coefficients in the Region of Interest (ROI), E_{ROI} is the set of expandable coefficients in the ROI, and BW is a binary , was previously determined using equation.

- Finally, coefficients adjacent to the boundaries between ROI and RONI that are left unchanged to avoid cross-region interference during embedding are known as non-changeable coefficients (N_{ROI}, N_{RONI}).

Step 3: Capacity Calculation

- The embedding capacity in RONI accounts for two bits per expandable coefficient:

$$\text{Capacity}_{RONI} = 2 \times |E_{RONI}| \tag{4.14}$$

- The capacity in ROI considers one bit per expandable coefficient:

$$\text{Capacity}_{ROI} = |E_{ROI}| \tag{4.15}$$

- Total embedding capacity available in the image is:

$$\text{Capacity}_{total} = 2 \times |E_{RONI}| + |E_{ROI}| \tag{4.16}$$

4.2.9 Adaptive Capacity Thresholds

The thresholds for embedding the watermark, denoted by T for the region of interest (ROI) and T_2 for the region outside the region of interest (RONI), are dynamically adjusted according to the required watermark size. Specifically, if the watermark length allocated for the ROI exceeds the absolute embedding capacity of the ROI, the threshold T is increased by one. Similarly, if the watermark length allocated for the RONI is greater than twice the absolute embedding capacity of the RONI, the threshold T2 is increased by one. as per the iteration strategy discussed in [Nuha], This iterative adjustment continues until the embedding capacity suffices to accommodate the watermark payload.

This capacity-adaptive classification achieves balanced watermark embedding, maintaining the quality of the ROI, and takes advantage of the larger space available in RONI, achieving high embedding rates with guaranteed reversibility.

4.2.10 Adaptive Embedding Process

To embed the watermarks generated through the watermark creation process, which are incorporated into the merged parameter matrix, the next steps are followed:

1. For each RONI coefficient, the prediction error e is classified into two categories based on a predefined threshold T. If the prediction error e is less than or equal to T, the watermarked value is determined by the following formula:

$$I^w(x,y) = \begin{cases} I'(x,y) + 4e + b, & if\ I'(x,y) \leq I(x,y), \\ I'(x,y) - 4e - b, & otherwise. \end{cases} \tag{4.17}$$

where b ∈ represents two bits.

For each RONI pixel, if the prediction error e is greater than the predefined threshold T', the pixel is classified as shiftable. In this case, the prediction-error is shifted by δ'

without embedding any data. The watermarked value is then calculated using the following formula:

$$I^w(x,y) = \begin{cases} I'(x,y) + \delta', & if\ I'(x,y) \leq I(x,y)' \\ I'(x,y) - \delta', & otherwise. \end{cases} \quad (4.18)$$

where δ'= 3T'+3.

2. For each ROI pixel, the prediction error e is compared to a predefined threshold T. Based on this comparison, e is categorized into two cases:

   1) If e ≤ T, the pixel is considered an expandable pixel of ROI. In this case, the watermarked pixel is computed using the following formula:

$$I^w(x,y) = \begin{cases} I'(x,y) + 2e + b, & if\ I'(x,y) \leq I(x,y)' \\ I'(x,y) - 2e - b, & otherwise. \end{cases} \quad (4.19)$$

   where b ∈ represents one bit of embedded data.

   2) If e > T, the coefficients is considered a shiftable coefficients of ROI. In this case, no data is embedded, and the watermarked coefficients is calculated by shifting the original pixel by δ, where δ = T + 1.

3) $$I^w(x,y) = \begin{cases} I'(x,y) + \delta, & if\ I'(x,y) \leq I(x,y)' \\ I'(x,y) - \delta, & otherwise. \end{cases} \quad (4.20)$$

This categorization allows the embedding of one bit in each expandable ROI coefficients while preserving image quality by avoiding embedding in coefficients with large prediction errors.

4.2.11 Watermarked Image

Following the embedding process, the watermarked image is obtained by performing reversing the steps in the section (4.2.7):

1. the Combined transformation components carrying the watermark are separated to obtain the four sub-bands of the image, namely the approximation coefficients and the horizontal,

vertical, and diagonal detail coefficients, while preserving the hierarchical structure derived from the wavelet transform.

$$[CA, CH, CV, CD] = sub - bands_{merged} \quad (4.21)$$

2. the inverse transform is performed using the same integer wavelet transform (IIWT) to reconstruct the final watermarked (stego) image.

$$[\,LL \;\; LH \;\; HL \;\; HH\,] \overset{IIWT}{\rightarrow} C^w(x, y) \quad (4.22)$$

Where $[\,LL \;\; LH \;\; HL \;\; HH] = I^w(x, y)$

4.2.12 Key Generation

To increase the security of the system, a secret key is generated at the conclusion of the embedding process. This key is constructed by concatenating the binary values of the following parameters. The security of the scheme is enhanced through the use of this key, ensuring that watermark extraction and access to patient data are limited exclusively to authorized entities.

- T and T_2: Threshold values, each encoded with 8 bits.
- L_1 and L_2: Lengths of the two watermarks, each stored with 20 bits.
- V: The number of polygon vertices that define the (ROI) in the image, stored with 10 bits.
- vertex: The coordinates of each vertex forming the polygon that defines the ROI.
- EPR_{len}: The length of the (EPR).
- V_{CA}: The number of polygon vertices at the CA sub-band level, stored with 10 bits.
- vertex_{CA}: The coordinates of each vertex at the CA sub-band level.

4.2.13 Extraction Process

The system requires two primary inputs for extraction and recovery the watermarked image and the secret key:

- Original secret key: The same auxiliary information generated after the embedding process is retrieved from the key, as mentioned in Section (4.2.12).

- The lifting scheme for IWT is reapplied, the watermarked medical image is decomposing to get the four frequency sub-bands, these subcomponents are combined just as in the embedding process.
- Based on the vertex coordinates extracted from the secret key, the watermarked medical image is segmented into two regions: the ROI, which is accurately separated based on the extracted coordinates, and the RONI, which includes all remaining areas of the image.
- Generate binary masks of the ROI on the approximation sub band and the full image using polygon vertices that are retrieved from the secret key.
- The pixel values in the first row and first column are copied directly from the watermarked image without any modification in the extraction; because they were not changed through embedding.

$$I_r(i,j) = \begin{cases} I_w(i,j), & \text{IF } i = 1,\ j = 1 \\ \text{(restored by embedding method in the methodology)}, & \text{Otherwise} \end{cases} \quad (4.24)$$

- Starting from the second row and second column, the predicted value is calculated at each the merged matrix value obtained from equation (44), using the same procedure used for embedding.

$$\hat{I}_r(x,y) = \left\lfloor \frac{I_r(x-1,y) + I_r(x,y-1)}{2} \right\rfloor \quad (4.25)$$

The prediction error is calculated:

$$e_r(x,y) = I^w(x,y) - \hat{I}_r(x,y) \quad (4.26)$$

- The values of the combined matrix are classified into groups based on the prediction error.

1. For each RONI values, the prediction error e_r is categorized into two cases based on a predefined threshold T'. If e_r is less than or equal to 4T' + 3, the pixel is considered expandable. In this case, the embedded data is extracted as follows:

$$b = |e_r(x,y)| - 4\left\lfloor \frac{|e_r(x,y)|}{4} \right\rfloor, \tag{4.27}$$

The original pixel value is then calculated by:

$$I(x,y) = \begin{cases} I'(x,y) + \lfloor e_r \,/_4 \rfloor, & if\ I'(x,y) \le I^w(x,y), \\ I'(x,y) - \lfloor e_r \,/_4 \rfloor, & otherwise. \end{cases} \tag{4.28}$$

If e_r is greater than 4T' + 3, the pixel is considered shiftable, and no data is extracted. The original pixel value is calculated by:

$$I(x,y) = \begin{cases} I'(x,y) + \delta', & if\ I'(x,y) \le I^w(x,y), \\ I'(x,y) - \delta', & otherwise. \end{cases} \tag{4.29}$$

where δ’ = 3T’ + 3.

2. For each ROI pixel, the prediction error e_r is classified into two conditions based on a predefined threshold T. If e^w is less than or equal to 2T+1, the pixel is considered expandable. In this case, the embedded data is extracted as follows:

$$b = e^w - 2\lfloor e_r \,/_2 \rfloor \tag{4.30}$$

The original pixel value is then calculated by:

$$I(x,y) = \begin{cases} I'(x,y) + \lfloor e_r \,/_2 \rfloor, & if\ I'(x,y) \le I^w(x,y), \\ I'(x,y) - \lfloor e_r \,/_2 \rfloor, & otherwise. \end{cases} \tag{4.31}$$

If e_r is greater than 2T+1, the pixel is considered shiftable, and no data is extracted. The original pixel value is calculated by:

$$I(x,y) = \begin{cases} I'(x,y) + \delta', & if\ I'(x,y) \le I^w(x,y), \\ I'(x,y) - \delta', & otherwise. \end{cases} \tag{4.32}$$

where δ = T + 1.

- The output of the extraction process consists of the extracted watermark bits embedded in the RONI and ROI regions, respectively, and the recovered merged matrix of the range coefficients.

$$[\text{bits}_{RONI}, \text{bits}_{ROI}, subband_{merged}]$$

- Inverse Wavelet Transform to Recover Image

The recovered merged sub bands matrix is split back into four sub bands with Same its sizes as before merging:

$$CA_{ex}, CH_{ex}, CV_{ex}, CD_{ex}$$

the spatial domain image is reconstructed by the inverse IWT for these sub bands:

- To restore the pixel values of the image to their original state before applying the adaptive shifts, we remove the shifts we performed in Sections 1, 2, and 2.2, which were applied to the image before the embedding process. This is done by subtracting the shift values added to both the ROI and ROI.

The final recovered image is obtained as follows:

$$Img(i,j) = \begin{cases} Img(i,j) - \Delta_{\text{ROI}}, & \text{if } mask(i,j) \neq 0 \text{ (pixel in ROI)} \\ Img(i,j) - \Delta_{\text{RONI}}, & \text{if } mask(i,j) = 0 \text{ (pixel in RONI)} \end{cases} \quad (4.33)$$

4.2.14 Hash Calculation and Watermark Integrity Verification

The SHA-256 hash values of the ROI and RONI regions of the extracted medical image are calculated and compared with the hash values extracted from both watermarks to ensure the integrity of the extracted image.

4.2.15 Tamper Detection, Localization and Recovery Process

If the integrity of the region of no interest (RONI) is verified, and the integrity of the region of interest (ROI) is not verified, a systematic approach is followed to detect tampering in medical images, accurately identify tampered regions, and restore the image content. The approach consists of the following sequential steps:

- To enable block-level tampering detection, the region of interest (ROI) of the recovered medical image is divided into non-overlapping 16x16 pixel blocks.

- The hash value of each block is calculated using the SHA-256 algorithm, and the hash value is reduced using the same process used during the watermark embedding phase.
- The hash values of the blocks extracted from the recovered image are compared to the corresponding original hash values extracted from the first watermark. Differences indicate tampering in these blocks.
- The pixels of the tampered blocks are highlighted by adjusting their intensity to the maximum possible value to clearly highlight the tampered regions for users, making it easier to identify tampered areas.
- Pixels identified as tampered with are recovered by replacing their values with values from a compressed version of the original region of interest extracted from the initial watermark data, thus restoring the original image information.

4.3 Performance Measures for Test and Evaluation

The performance of the proposed method will be evaluated using key image quality measures, including perceptual similarity, embedding capacity, reversibility, and robustness. These metrics serve as important indicators of the method's effectiveness in embedding data while preserving the visual quality and integrity of the host image. These metrics will provide a comprehensive assessment of the method's capabilities and allow for objective quantification and comparison with existing reversible watermarking approaches.

1. Peak Signal to Noise Ratio (PSNR):

   (PSNR) is used to evaluate the imperceptibility of the watermarked image. It is calculated using the following formula:

$$PSNR = 10\, log_{10} \frac{(I_{max})^2}{MSE} \mathrm{d}B \qquad (4.35)$$

   where I_{max} is the maximum pixel value in the image and MSE is the mean squared error between the watermarked image and the original image.

2. Structural Similarity (SSIM):

The structural similarity (SSIM) index will be used to evaluate the visual quality of the watermarked images. SSIM will measure the similarity between the original and watermarked images, with values ranging from 0 (no similarity) to 1 (identical). The high SSIM values, all close to 1, indicate that the proposed watermarking technique is able to embed data with minimal perceptual distortion. It is defined as:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_1)}. \qquad (4.36)$$

where μ_x is the mean value of the region x, and μ_y is the mean value of the region y .while C1 and C2 are the small constants, μ_x is the mean value of and μ_y is the mean value of y . and σ_x^2 is the variance of x, and σ_{xy} is the covariance of x and y [63].

3. Embedding capacity:

The embedding capacity, or the amount of data that can be reversibly embedded, is an important performance metric. The proposed method aims to achieve high embedding capacities while maintaining high visual quality. Embedding capacity, measured by the total payload in bits and the embedding ratio in bits per pixel (bpp), is a crucial performance metric for evaluating the effectiveness of the proposed method [64]. The embedding capacity of the proposed watermarking system will be calculated From the following equation, expressed in bits per pixel (bpp) [65].

Capacity = total number of required pixels for embedding / total host image size (16).

4. Reversibility and Perceptual Quality:

The proposed reversible watermarking method aims to achieve perfect reversibility, allowing lossless reconstruction of the original image after data extraction. The perceptual quality of the restored image will be verified through objective metrics such as (PSNR) and (SSIM) [66], with expected results of infinite PSNR (no distortion) and SSIM of 1 (identical to original), validating the method's effectiveness in fully recovering the host image without any degradation, a crucial requirement for sensitive applications.

5. Image Fidelity

Image fidelity measures the similarity between the original image I and the watermarked image $\hat{I}$ [13]:

$$\text{Fidelity} = 1 - \frac{\sum\ (I(x,y) - \hat{I}(x,y))^2}{\sum\ I(x,y)^2} \qquad (4.37)$$

Higher values (closer to 1) indicate better preservation of image quality.

6. Average Difference

Average difference quantifies the mean absolute pixel differences, this standard metric complements PSNR/SSIM for comprehensive quality assessment[45] :

$$\text{Average Difference} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} |I(i,j) - \hat{I}(i,j)| \qquad (4.38)$$

Lower values indicate minimal visual distortion between original and watermarked images.

7. Robustness Evaluation:

The proposed method aims to ensure the authenticity and integrity of DICOM medical images by embedding a watermark that can be retrieved and matched exactly with the original if and only if the pixel data and header information remain unaltered. However, any manipulations of the image data, whether intentional or unintentional, can corrupt the embedded watermark, resulting in a mismatch between the original and retrieved watermarks. To assess the robustness of the system, various image processing operations will be applied, including malicious manipulations such as adding or removing parts of the image. to evaluate the robustness of watermarking schemes, bit error rate (BER) and normalized cross correlation (NCC) between original watermarks and extracted watermarks are calculated respectively according to Equations (4.39) and (4.40). The value of BER is closer to 0 and the value of NCC is closer to 1, the robustness of watermarking scheme is stronger. Common attacks with different parameters, are applied to test the watermarking robustness. like impulse noise (salt-and-pepper), Gaussian noise, cropping, and collage attacks

original and extracted watermarks are calculated respectively according to Equations (4.39) and (4.40). where will be used to evaluate the robustness, with BER closer to 0 and NCC closer to 1 indicating stronger robustness. Common attacks such as cropping, salt-and-pepper noise, and other distortions will be used to test the watermarking scheme's resilience.

$$BER = \frac{Number\ of\ incorrectly\ decoded\ bits}{Total\ number\ of\ bits} \tag{4.39}$$

$$NCC = \frac{\sum_{i=1}^{H} \sum_{j=1}^{W} (W_o(\mathrm{i}, j) \times W_e(\mathrm{i}, j))}{\sum_{i=1}^{H} \sum_{j=1}^{W} W_o^2(\mathrm{i}, j)} \tag{4.40}$$

Where H and W are the height and the width of watermark, W_o is the original watermark, W_e is the extracted watermark, i and j are coordinates of pixels in these watermarks [56].

4.4 Summary

In this chapter the scheme of the proposed reversible watermarking scheme is explained, where the image is segmented into regions of interest (ROI) and regions of non-interest (RONI). The image is decomposed into subdomains for embedding and extraction using the Integer Wavelet Transform (IWT). Prediction error expansion is used to embed watermarks consisting of patient data, hash values, and tampered data recovery information. The embedding is adaptively performed to balance capacity and image quality. An extraction process is applied to retrieve both the embedded data and the medical image, and the integrity of both is verified by comparing the hash values. Using the embedded ROI data, tampered blocks are identified and recovered. This approach achieves secure embedding, accurate extraction, tamper detection, and lossless data recovery.

Chapter 5

Experimental Results and Discussions

5.1 Introduction

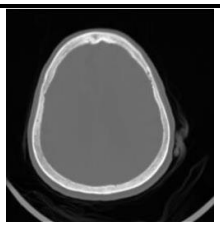
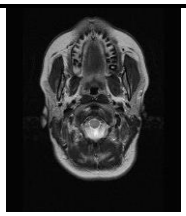
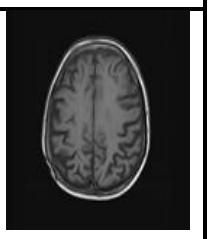
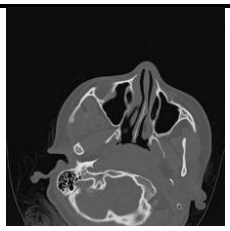
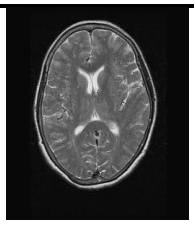
This chapter presents the performance analysis and experimental results of the proposed reversible watermarking methodology for protecting medical DICOM images, the details of which were explained in previous chapters. Several tests were conducted to evaluate the performance of the proposed method and compare its performance with existing methods. We provide a description of the data and tools used, followed by the presentation of results of the proposed method using various computational benchmarks. Finally, the performance is evaluated in comparison with other methodologies.

5.2 Experimental results and analysis

5.2.1 Experimental Process:

To verify the performance of the proposed algorithm, a series of tests were conducted using the MATLAB R2020a simulation environment on a personal computer equipped with an Intel Core i7 processor 10th Generation. The images used included 40 clinical brain images in DICOM format, including both magnetic resonance imaging (MRI) and computed tomography (CT) scans. These images were obtained from the archives of Zawia Teaching Hospital, Zawia, Libya and featured a variety of technical characteristics. They were all in grayscale format, with different pixel dimensions (768×768, 512×512, and 256×256), and with bit depths ranging from 12 to 16 bits, reflecting the true diversity of clinical medical imaging data. Table 5.1 shows the specifications of a subset of host images from the DICOM medical image reverse watermarking experiment dataset.

Table 5-1 Representative Subset of the Host DICOM Images

Filename	CTimg1		**Filename**	MRIimg11	
Dimensions	512 * 512		**Dimensions**	512 * 512	
Modality	CT		**Modality**	MRI	
Bit depth	12		**Bit depth**	12	
Filename	CTimg34		**Filename**	MRIimg17	
Dimensions	512 * 512		**Dimensions**	256 * 256	
Modality	CT		**Modality**	MRI	
Bit depth	12		**Bit depth**	12	
Filename	CTimg23		**Filename**	MRIimg37	
Dimensions	768 * 768		**Dimensions**	512 * 512	
Modality	CT		**Modality**	MRI	
Bit depth	12		**Bit depth**	16	

The brain image segmentation process is performed to separate the medical image into the (RONI) black background, which contains no vital diagnostic data, and the region of interest (ROI), which holds important diagnostic information for the physician. The adequacy of K-means for isolating the head region is confirmed in the experimental results, where it successfully segmented ROI areas comprising 26-50% of the total area size across all test images. Figure 8 illustrates the ROI and RONI regions for two sample images (CT and MRI). Figure 9 shows the original image of the brain and its tissue segmentation.

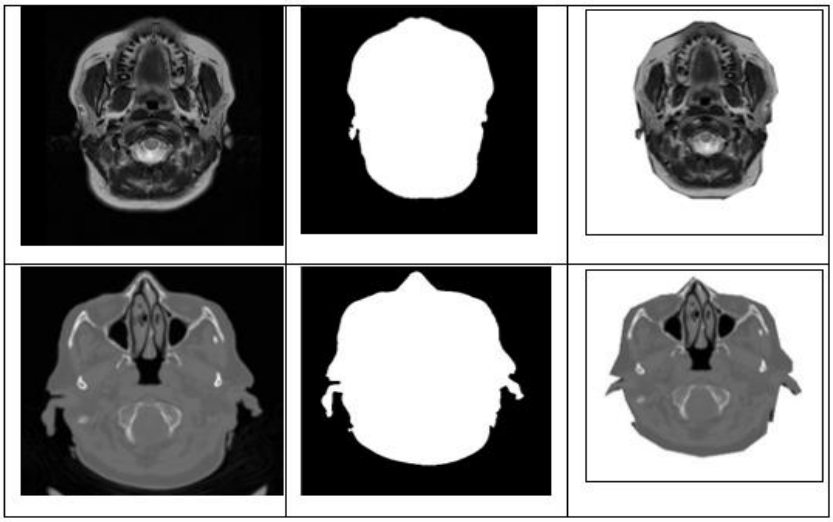


Figure 8 ROI, and RONI for an MRI, CT image

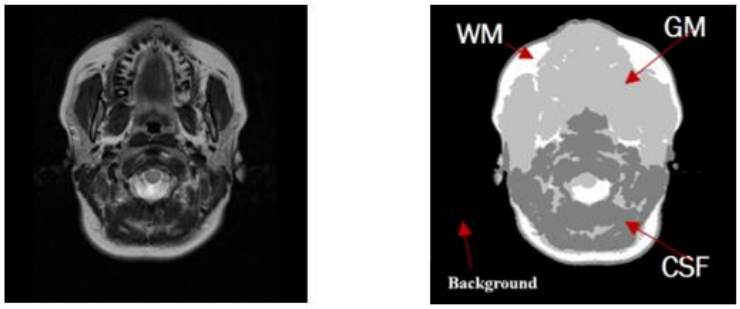


Figure 9 Brain tissue segmentation

5.2.2 Visual Quality and Capacity Results

From the results presented in Table 5.2, we observed a balance between watermark capacity and image quality for both the diagnostic (ROI) and the background (RONI). The results demonstrate a clear relationship between embedding rates and image quality preservation, with PSNR values ranging from 41.90 to 62.02 dB, and SSIM scores ranging from 0.7767 to 0.9946. This is consistent with what is explained in [49], that to ensure the human eye cannot perceive the embedded information, the PSNR value must exceed 36 dB. This indicates that image integrity was preserved despite the high amplitude. PSNR values were higher in RONI regions (between 46.71 and 70.05 dB), reflecting the ability of the system to adapt to the characteristics of each region, and that the nature of the RONI region in the frequency domain is well-homogenized, enabling effective embedding. Moreover, SSIM values consistently remain above 0.95 in most cases, demonstrating excellent structural preservation. In some images, where embedding rates approach the system's capacity limits, quality metrics may be relatively lower. However, the results demonstrate acceptable quality. The table shows remarkably high capacity values, up to 0.903 bpp, while maintaining excellent visual quality. This indicates space savings in storing the location map data, along with efficient utilization of available bits and frequency coefficients for effective embedding of watermark data. These results demonstrate the effectiveness of the proposed inverse watermarking system in balancing payload capacity and visual quality preservation in medical imaging applications.

Table 5-2 Capacity and Visual Quality of Watermarked Images

Image	ROI Size%	ROI Capacity (bit)	ROI Capacity (bpp)	PSNR (ROI)	RONI Capacity (bit)	RONI Capacity (bpp)	PSNR (RONI)	Total Capacity (bit)	Capacity (bpp)	Total PSNR	SSIM
CTimg1	43.62	25267	0.1740	48.82	172818	0.5348	66.51	198085	0.7089	51.26	0.9807
CTimg34	26.9	28600	0.3212	45.47	100268	0.1594	51.14	128868	0.4806	46.42	0.9509
CTimg23	43.93	12395	0.0478	51.41	21706	0.0656	53.78	34101	0.0578	52.30	0.9839
MRIimg11	37.27	14444	0.1131	40.50	207532	0.7333	46.37	221976	0.8465	41.90	0.7767
MRIimg17	31.93	11450	0.4210	45.26	45674	0.4295	46.71	57124	0.8506	45.67	0.8579
MRIimg37	39.65	13153	0.0971	60.11	232094	0.7997	70.05	245247	0.8969	62.02	0.9946

5.2.3 Reversibility Results

The proposed scheme was verified to be completely reversible for the full medical image, not just the region of interest (ROI). To achieve complete and accurate recovery of the original image without any loss of information after watermark extraction, the quality of the recovered image was evaluated using several metrics. The results of these tests for the image samples are presented in Table 5.3, where the Peak Signal-to-Noise Ratio (PSNR) reached infinite values, indicating complete recovery of the original image without any quality loss. The structural similarity index (SSIM), normalized cross-correlation coefficient (NCC), and image fidelity all reached the ideal value of 1, while the Average difference was zero, confirming that the recovered image perfectly matches the original. Furthermore, the EPR extraction rate achieved 100%, demonstrating the complete retrieval of embedded information without any loss. These results confirm the ability of the novel to maintain the integrity of the original image and the watermark completely after the embedding and extraction process.

Table 5-3 Evaluations for Original and Recovered Cover Image

Performance evaluations for original and recovered cover image						
Image Name	**PSNR**	**SSIM**	**NCC**	**Image fidelity**	**Average difference**	**EPR Extracted%**
CTimg1	inf	1	1	1	0	100%
CTimg36	inf	1	1	1	0	100%
CTimg23	inf	1	1	1	0	100%
MRIimg11	inf	1	1	1	0	100%
MRIimg17	inf	1	1	1	0	100%
MRIimg37	inf	1	1	1	0	100%

5.2.4 Evaluating the tampering and robustness of the proposed scheme.

To ensure accurate clinical diagnosis, the integrity and accuracy of the Region of Interest (ROI) must be maintained. To detect any manipulation of the ROI, the proposed method, when the non-ROI is intact, uses watermark authentication to verify its integrity by comparing the block hash values of the ROI in the extracted image with those of the original image. The results

showed a perfect match when the ROI was unaltered. If the ROI was manipulated, it could be recovered using the watermark recovery.

Through the testing phase, code was employed to simulate a variety of attacks aimed at manipulating the watermark image. This included implementing manipulation detection and recovery attacks, such as clipping, clustering, and shearing. Gaussian and salt and pepper noise were also implemented to verify the system's robustness, privacy protection, and attack resistance.

The results of all these attacks were evaluated after watermark extraction using various image quality indicators, such as PSNR and SSIM, along with the EPR extraction rate, to ensure a comprehensive and accurate assessment of the system's performance in tamper detection and recovery of damaged content, as well as in noise resistance and privacy protection.

5.2.5 Tamper localization and image Recovery:

The proposed system will be evaluated under intentional attacks that aim to alter or remove parts of the image. This test assesses the efficiency of the system in detecting intentional modifications and restoring the original content. with providing a detailed analysis of each attack type (shear, collage, and cropping).

Table 5.4 presents the results of the proposed method for detection, localization and recovering manipulated blocks after applying a shearing attack with size 100×100 pixels to evaluate resistance to geometric distortion, with a distortion ratio of 0.5. The results indicate that the proposed scheme achieves accurate identification of tampered locations within the region of interest (ROI), with highly efficient recovery. A little variation in the metric of recovery quality was observed between images, depending on the characteristics of the image and the tampered region. However, the overall values for the recovery quality metric and the watermark extraction rate were high. The structure of the recovered image was fully preserved without any loss of structural information. This achieves the objective of the novel in ensuring accurate and reliable recovery.

Table 5-4 Shear Attacks Results

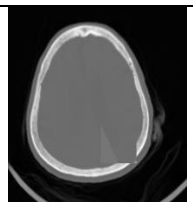
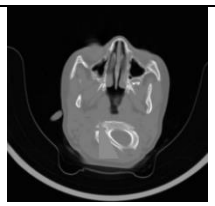
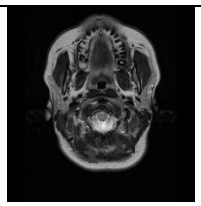
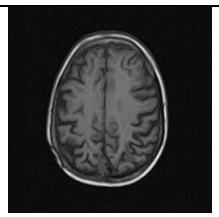
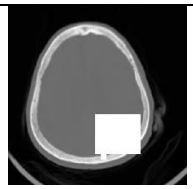
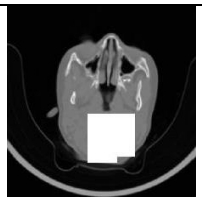
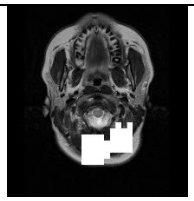
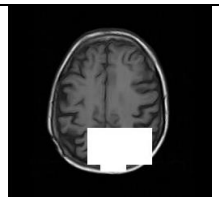
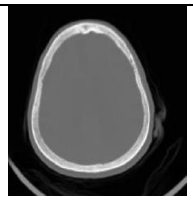
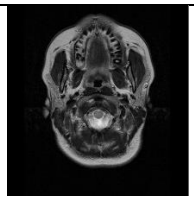
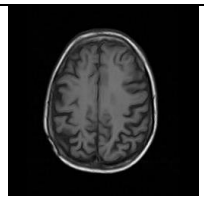
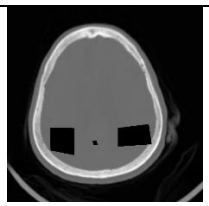
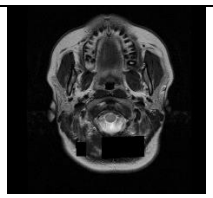
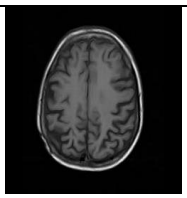
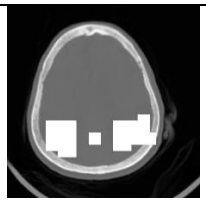
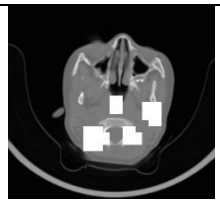
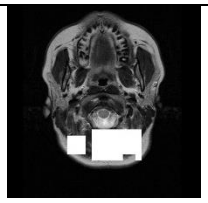
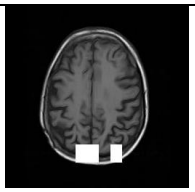
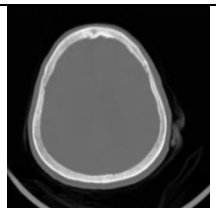
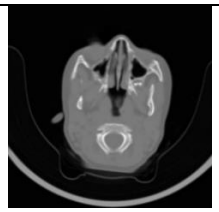
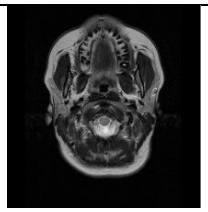
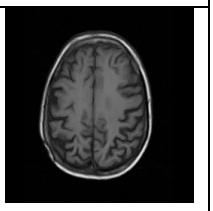
Image Name	CTimg1	CTimg34	MRIimg11	MRIimg17
Tampered Images (100 *100) pixel				
Tamper localization				
Recovered image				
PSNR	102.953984	97.125124	98.673859	100.131685
SIIM	1	1	1	1
EPR Extracted%	100%	100%	100%	92.24%

Table 5.5 presents the simulated cropping attack results for four test images. This attack removes an actual part of the image and deletes it completely. This tamper was applied to various sizes exceeding 100x100 pixels. The results display the localization of tampered regions, and the images after the tamper was recovered. PSNR values range from 98.22 to 103.63, indicating that the recovered medical images achieved high quality after recovery. The proposed approach maintained the structural integrity of the recovered images, with the SIIM score registering a value of 1 for all tested images. Furthermore, the extracted watermarking from the tested images reached 100%, demonstrating the robustness of system and its ability to recover embedded medical information without any loss despite cropping.

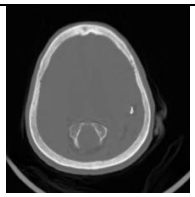
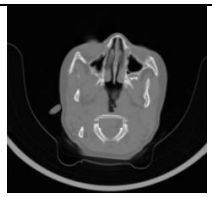
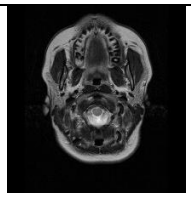
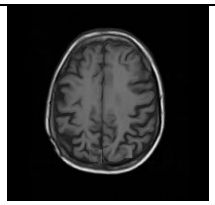
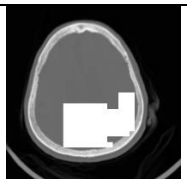
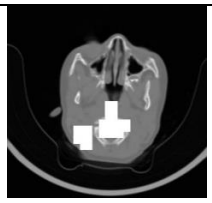
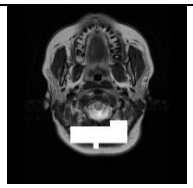
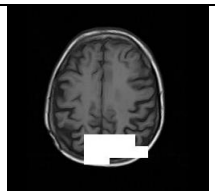
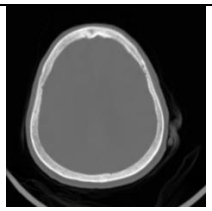
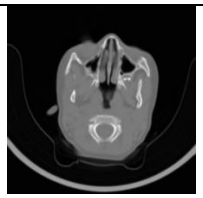
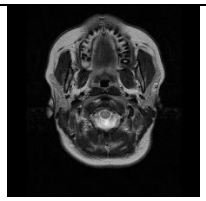
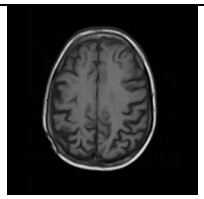
Table 5-5 Results of the Cropping Attacks

Image Name	CTimg1	CTimg34	MRIimg11	MRIimg17
Tampered Images				
Tamper localization				
Recovered image				
PSNR	102.965346	99.063554	98.220580	103.639576
SIIM	1	1	1	1
EPR Extracted%	100%	100%	100%	100%

In Table 5.6, the results of collage attacks are presented, where a specific region with different dimensions exceeding (100 × 100) is copied and pasted to another region in the image. This attack aims to test the system's ability to detect copying and pasting operations within an image. to test the system's ability to detect copying and pasting operations within an image. These results show the efficiency of the proposed scheme in detecting, locating, and recovering tampered brain images. The values of PSNR were from 99.62 to 101.39; this confirms the high quality of the recovered images. The SIIM structural integrity index, which achieved the ideal value of 1 in all samples, indicates that the image structure remained intact without distortion. The watermark recovery rate was 100% in the 512×512 images, while it decreased little to 96.33% in the size (256×256) images. This is due to the slight decrease in the amount of detail available in the smaller images, which may affect data retrieval accuracy. These results show

the robustness and reliability of the scheme for collage attacks, while considering the impact of image size on enhancing recovery performance.

Table 5-6 Results of the Collage Attacks

Image Name	CTimg1	CTimg34	MRIimg11	MRIimg17
Tampered Images				
Tamper localization				
Recovered image				
PSNR	101.392592	99.991659	99.618534	102.186201
SIIM	1	1	1	1
EPR Extracted%	100%	100%	100%	96.33%

5.2.6 Evaluating Robustness Against Attacks

In addition to deliberate attacks, the proposed approach was also found to be effective against unintended attacks such as salt and pepper noise and Gaussian noise.

In this scheme, a Gaussian noise attack is simulated as mean 0 and various values of variance. Table 5.7 presents the performance of the proposed system under Gaussian noise attack. This result demonstrates the balance between watermark robustness and visual quality. The system maintained the integrity of the watermark up to a variance of 0.010, with a low bit error rate of 0.000050. The system exhibited a gradual decline in visual quality from NCC = 1.0000 to NCC = 0.8608 at 0.025, reflecting the ability of the scheme to achieve acceptable visual quality even

at 0.025 noise. A shown in figure 10 the bit error rate (BER) stabilized around 50% starting at 0.020 noise, meaning that approximately 50% of the watermark could be correctly extracted even at noise levels exceeding 0.025.

Table 5-7 performance results under Gaussian noise attack

Parameter	NCC	SSIM	BER
0.000	1.0000	1.000000	0.000000
0.001	0.9932	1.000000	0.000000
0.010	0.9398	0.999000	0.000050
0.020	0.8852	0.998000	0.425204
0.025	0.8608	0.997500	0.492104

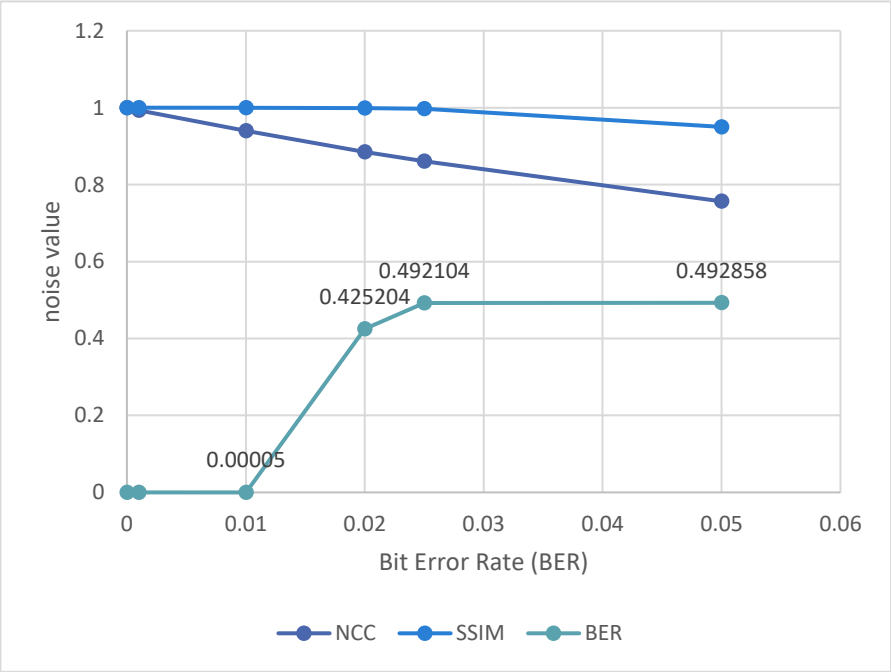


Figure 10 BER of watermark under Gaussian noise

5.2.7 COMPARISONS WITH ANOTHER SCHEME

Table 5-8 Comparison of Proposed Method with Recent Techniques

Scheme	Purpose	methodology			Capabilities						Performance	
		Embed Method	Embed Domain	ROI-Based	EPR Hiding	DICOM Support	Integrity Mechanism	Tamper Localize	Tamper Recovery	Reversible (No attack)	Capacity (bpp)	PSNR SSIM
[44] Abokhdair (2018)	Data hiding - integrity	Adaptive PEE	Spatial	✓ Auto K-means + line simplify	✓ (Patient data in ROI)	✓ (MRI/CT DICOM)	MD5 per region block average block-level	Block-level (via block average)	ROI only (recovery data in RONI)	✓ (Full image)	≈1.5 max	40-77dB ≈68.62dB SSIM≈0.9586
[20] Pal (2023)	Data hiding - authenticity	Rule-based coefficient adjust	Frequency (LWT)	Auto SVM-based NROI)	✗ (logo 165×165) (Cover 256×256)	NR	✗ Watermark based	✗ (no recovery)	✗	Claimed reversible (SSIM<1)	≈ 0.9 bpp 63516 bit	≈66–67 dB SSIM≈ 0.9694
[46] Bhalera (2023)	Data hiding - integrity	PEE (ROI) + Diff. Hist. Shift (RONI)	Spatial	✓ manual fixed ROI 100×100 pixel	✗	NR	SHA-1 (ROI blocks) + RONI hash	Block-level (ROI)	ROI only (compressed ROI in RONI)	✓ (Full image)	0.15 bpp 160-bit ROI 45360-bit RONI	≈55 dB / SSIM≈0.99 (512×512 cover)
[47] Arham (2023)	Data hiding - integrity	Multi-layer DE (ROI) + 2D-DE (RONI) + LSB (border)	Hybrid	✓ Manual polygon ROI (4–12%)	✓ (Patient records in ROI)	NR	MD5 + block-average	Block-level ROI/RONI	Full image (ROI via copy in RONI)	✓ RONI-ROI ✗ Boundary	ROI ≈ 6 RONI ≈ 0.8, border = 1	Up to ≈ 50.86 dB, SSIM≈0.99
[48] Shi (2021)	Data hiding - integrity	Modified DE (ROI) + Histogram mod RONI	Spatial	✓ semi-Auto (active contour)	✓ (Privacy watermark)	NR	Parallel lattice hash	Block/ Region two- level	ROI-only approximate recovery	✓ SSIM=1 `	Up to ≈1.0 bpp	60.6 dB
[49] Hurrah (2020)	Data hiding - authenticity	LSB + mean modification	Spatial	✗ (full image as carrier)	✓ (EMR, 16 KB)	NR	Fragile mean modification	Block-level	✗ (no recovery)	✗ (semi) PSNR≈58.50 SSIM≈0.9913	(16 KB / 256×256 512×512 cover)	>47 dB
[50] Geetha (2020)	Data hiding - integrity	Rhombus mean interpolation	Spatial	✗	✓	NR	Checksum per block	Block-level	✗ (no recovery)	✓	1.5 bpp (393,216 bits) 512×512 cover)	42.40 dB SSIM 0.9821
[51]Li (2024)	Content authenticity	LSB + Chaotic Mapping	Frequency (DWT)	✓ Segment blocks	✗	NR	average higher 4 bits of block	Block-level	✓ Approximate	✓	NR	43.8 dB

Scheme	Purpose	methodology			Capabilities						Performance	
		Embed Method	Embed Domain	ROI-Based	EPR Hiding	DICOM Support	Integrity Mechanism	Tamper Localize	Tamper Recovery	Reversible (No attack)	Capacity (bpp)	PSNR SSIM
[52] Nguyen (2021)	authenticity	PEE	Spatial	✗	✗	NR	Block-wise authenticity code	Block-level	✗	✓	avg ≈25,191 bits	≈48–52 dB
[54] Lin (2022)	EPI hiding	ENMI interpolation (2×2→3×3)	Spatial	✗ (uniform blocks)	✗ (EPI only, no full EPR)	✗	authenticity bit via mod 2^ Threshold	Block-level (3×3)	✗	✓ (via corner pixels)	0.55bpp	≈46.94 dB SSIM≈ 0.984
[55] Awasthi (2023)	Data hiding-authenticity	Schur + LWT-DCT-SVD	frequency (3-level LWT)	✗ (full image as carrier)	✗ (logo + patient text)	✓ (DICOM ultrasound)	SURF features for authentication	✗	✗ no recovery	✗	NR (two 256×256 WMs)	>40 dB no attack
[57] Bamal (2024)	Data hiding - integrity	SLT + RS vector + ANN	Hybrid (SLT + spatial)	✓ ANN-detected ROI	✗ (biometric ID) patient ID	NR	SHA-3 + AES + LZW	block (ROI)	ROI only (small diagnostic area)	✓ (ROI) ✗ (full image due to preprocessing)	1.1225 bpp (62K bits) 512×512 cover	≈ 51.62 dB
Proposed	Data hiding - authenticity integrity	LIWT + adaptive PEE (1-bit ROI, 2-bit RONI) + AES	frequency (LIWT	✓ Auto ROI ≈26–50% of image)	✓ Full EPR + DICOM metadata	✓ (Brain MRI - CT DICOM)	SHA-256 for ROI (block-level) ROI & RONI (region-level)	Block-level in (ROI) + region-level	recovery of ROI (data in RONI)	✓ Fully reversible 256×256 512×512 768 × 768 cover)	Up to ≈ 1 bpp 452245 bits/	Up to ≈70 dB / SSIM≈ 0.998

The analysis of the reviewed studies reveals divergent developmental paths in the field of medical image watermarking. It is described by parallel research approaches that combine advanced automation with manual methods. where Abokhdair et al. [53] in 2018 proposed fully automatic ROI segmentation using the k-means clustering algorithm. However, manual and semi-automatic approaches persisted in recent research, underscoring ongoing practical challenges, like in Bhalerao et al. [55] (2023), where the researchers applied a manual process to isolate a 100 x 100-pixel area of interest from the medical images by manually identifying the center point of this region, and in Arham et al. [56] (2023), which relies on manually selected polygon coordinates. This highlights the diversity of techniques used. Although automated methods for ROI segmentation exist, they face practical limitations such as computational cost and low embedding capacity, which constitute a major obstacle to fully utilizing the capabilities of these techniques, especially in reversible systems that provide ROI recovery in case of tampering. This conceptual development reflects both automated methods (as in [53], [13], [66]) and improved manual methods (as in [55], [56]), each serving specific purposes, which can be observed in the study by Bamal & Kasana [66] (2024). Although they used ANN to accurately determine ROI, they were practically limited to a small region with limited dimensions (30 × 100 neurons). Due to embedding capacity challenges, embedding diagnostic ROI recovery data requires high embedding capacity, which may degrade image quality or exceed the embedding capacity limits. Pal et al. [3] also pointed out the high computational complexity of the SVM model they implemented.

Table 5.8 systematically compares 12 state-of-the-art watermarking schemes against the proposed LIWT+adaptive PEE method based on clinical requirements. Pure spatial fragile systems excel in capacity but lack recovery: Hurrah embeds 16 KB of electronic medical records (reversible) yet offers no data restoration under tampering; Geetha achieves 1.5 bpp (393 Kbits) but limits to detection only; Lin provides 0.55 bpp with block-level detection (NCC=0.594@var=0.02) but no modified block recovery.

A robust method that is based on frequency loses reversibility: Awasthi maintains PSNR>30 dB post-attacks with DICOM compliance without metadata management but is non-reversible; Pal uses SVM for undefined NROI verification yet embeds only logos (no full EPR). Variance was revealed in the schemes of ROI: Shi offers approximate ROI-only recovery via

neighborhood differences; Bamal achieves precise ANN-detected ROI (1.12 bpp) but loses full reversibility due to preprocessing; Abokhdair provides complete recovery (PSNR=68 dB) capped at 1.5 bpp max.

The proposed method resolves these limitations through automatic ROI detection (26-50%), complete EPR/DICOM metadata embedding (patient records + full header integrity), lossless ROI recovery (∞ dB) from RONI, superior Gaussian robustness (NCC=0.885@var=0.02 vs. Lin's 0.594), and PSNR=70 dB/SSIM=0.998 at 1 bpp (452 Kbits), proving comprehensive clinical superiority.

It outperforms Abokhdair particularly in attack robustness and integrity verification: while Abokhdair relied on distortion-sensitive spatial PEE with MD5 region hashing (ROI/RONI block-averaging, collision-prone), the proposed method employs frequency-domain LIWT with adaptive PEE and SHA-256 block/region verification. Tests showed near-perfect NCC/SSIM≈1 under Gaussian noise up to var=0.01 (untested in Abokhdair), thanks to LIWT's preservation of low-frequency components less affected by added noise.

Table 5-9 Gaussian Noise Robustness Comparison

Variance	Evaluation Metric: NCC/ SSIM/ BER%						
	Proposed	**[48]**	**[54]**	**[49]**	**[50]**	**[55]**	**[57]**
0.0008	-	-	-	-	-	-	NCC<0.8
0.001	0.993 1.00 0.000%	-	-	-	-	NCC .999 SSIM 0.676	-
0.002	-	NCC 0.971	-	-	-	-	-
0.01	0.940 0.999 0.000%	-	-	-	-	-	-
0.02	0.885 0.998 42.52%	-	NCC 0.594 BER 40.64%	NCC 0.482 BER 61.46%	NCC 0.502 BER 49.86%	-	-
0.025	0.861 0.998 49.21%	-	-	-	-	-	-

Table 5.9 compares Gaussian noise performance at exact variance levels reported in literature. The proposed method maintains NCC ≥ 0.861 up to variance=0.025, outperforming Lin [63] (NCC=0.594 at 0.02), Hurrah [58] (0.482), and Geetha [59] (0.502) by 42-78% at equivalent noise levels.

5.3 Summary

To summarize thus far, the proposed LIWT-APEE reversible watermarking scheme was evaluated using 40 DICOM grayscale CT/MRI images. The experimental results demonstrate superior performance with PSNR of 40-70 dB, capacity up to 452 Kbits (1 bpp expandable coefficients in ROI + 2 bpp in RONI), infinite PSNR/SSIM/NCC recovery under zero attack, and Gaussian robustness (NCC=0.885@var=0.02 vs. Lin's 0.594). Additionally, the method enables tamper detection, localization, and recovery for ROI comprising 26-50% of the image. the proposed scheme achieved the superiority over other schemes Through comparative evaluation, in addressing research gaps and overcoming key challenges.

Chapter 6

Conclusion

7.1 Conclusion

This study proposes a reversible watermarking method that integrates Lifting-based Integer Wavelet Transform (LIWT) with adaptive Prediction Error Expansion (PEE) for brain DICOM images in MRI/CT for telemedicine applications. The method embeds patient records and SHA-256 hash values through automated Region of Interest (ROI) detection (26-50%), achieving 1 bit/pixel capacity in ROI + 2 bits/pixel in RONI (452 Kbits total), lossless recovery (∞ dB), superior imperceptibility (PSNR 40-70 dB, SSIM=0.998), precise tamper localization/recovery (>100×100 pixels), and Gaussian robustness (NCC=0.885@var=0.02 vs. Lin's 0.594), ensuring comprehensive clinical security for confidentiality, authenticity, and integrity. Experimental results validate the proposed algorithm's capability to guarantee patient record privacy and medical image authenticity/integrity. The results show high hiding capacity up to 1 bpp with low perceptual distortion, achieving PSNR up to 70 dB. We verify the full reversibility using several metrics, including recovered image quality with infinite PSNR, SSIM=1, and NCC=1. The algorithm enables block-level tamper detection and ROI recovery via SHA-256 hashes, with superior Gaussian robustness (NCC=0.885@var=0.02 vs. 0.594 for Lin distribution). Comparative analysis against recent medical watermarking schemes confirms the proposed method surpasses current limitations, addressing critical gaps in benchmarked studies by providing full DICOM compliance, HIPAA requirements, and metadata handling, thus offering a comprehensive solution for secure Electronic Medical Record transmission in clinical telemedicine applications.

7.2 Recommendation

The significance of this research lies in securing the exchange of digital medical images and telemedicine applications, thereby protecting patient privacy and ensuring the reliability and integrity of diagnostic data.

However, promising the results, several avenues for future research remain. Evaluating the algorithm's performance across diverse image modalities beyond CT/MRI and non-brain anatomies would enhance its versatility and applicability, given its current focus on brain CT/MRI formats. For a comprehensive robustness assessment, evaluations against other attacks are recommended, as this research provides a detailed assessment of Gaussian noise. To increase capacity, the embedding threshold could be optimized automatically using machine learning. In conclusion, this research presents a novel reversible watermarking algorithm for medical images that shows effective performance. It contributes significantly to medical image security and diagnostic reliability, with potential for safeguarding images in clinical telemedicine applications.

References

Uncategorized References

[1] K. Lata, C. Gupta, and L. R. Cenkeramaddi, "A cryptographic framework for secure medical imaging in smart healthcare environments," *Results in Engineering,* p. 106780, 2025.

[2] M. Magdy, K. M. Hosny, N. I. Ghali, and S. Ghoniemy, "Security of medical images for telemedicine: a systematic review," *Multimedia Tools and Applications,* vol. 81, no. 18, pp. 25101-25145, 2022.

[3] D. Mata-Mendoza, D. Nuñez-Ramirez, M. Cedillo-Hernandez, and M. Nakano-Miyatake, "An improved ROI-based reversible data hiding scheme completely separable applied to encrypted medical images," *Health and Technology,* vol. 11, no. 4, pp. 835-850, 2021.

[4] M. Eichelberg, K. Kleber, and M. Kämmerer, "Cybersecurity in PACS and medical imaging: an overview," *Journal of Digital Imaging,* vol. 33, no. 6, pp. 1527-1542, 2020.

[5] J. Ivanova *et al.*, "Regulation and Compliance in Telemedicine," *Journal of medical Internet research,* vol. 27, p. e53558, 2025.

[6] P. Sharma, S. P. Kiron, D. Aju, and A. M. Lal, "An enhanced security framework for secured transfer of DICOM images by ensuring the CIA triad," *Multimedia Tools and Applications,* pp. 1-26, 2025.

[7] S. Sankaranarayanan *et al.*, "Enhancing healthcare imaging security: color secret sharing protocol for the secure transmission of medical images," *IEEE Access,* vol. 12, pp. 100200-100216, 2024.

[8] S. Sharma, J. J. Zou, G. Fang, P. Shukla, and W. Cai, "A review of image watermarking for identity protection and verification," *Multimedia Tools and Applications,* vol. 83, no. 11, pp. 31829-31891, 2024.

[9] S. D. Mahmood, F. Drira, H. F. Mahdi, and A. M. Alimi, "Secure Medical Image Sharing: Technologies, Watermarking Insights, and Open Issues," *IEEE Access,* 2025.

[10] M. Larobina, "Thirty years of the DICOM standard," *Tomography,* vol. 9, no. 5, pp. 1829-1838, 2023.

[11] A. Al Siam, M. M. Hassan, M. A. Rahaman, and M. Abdullah, "Diegif: An efficient and secured DICOM to EGIF conversion framework for confidentiality in machine learning training," *Results in Control and Optimization,* vol. 18, p. 100515, 2025.

[12] G. Coatrieux, H. Maître, B. Sankur, Y. Rolland, and R. Collorec, "Relevance of watermarking in medical imaging," in *Proceedings 2000 IEEE EMBS international conference on information technology applications in biomedicine. ITAB-ITIS 2000. Joint Meeting Third IEEE EMBS international conference on information technol*, 2000: IEEE, pp. 250-255.

[13] S. Gull and S. A. Parah, "Advances in medical image watermarking: a state of the art review," *Multimedia Tools and Applications,* vol. 83, no. 1, pp. 1407-1447, 2024.

[14] A. Al-Haj, "Providing integrity, authenticity, and confidentiality for header and pixel data of DICOM images," *Journal of digital imaging,* vol. 28, pp. 179-187, 2015.

[15] A. F. Qasim, F. Meziane, and R. Aspin, "Digital watermarking: Applicability for developing trust in medical imaging workflows state of the art review," *Computer Science Review,* vol. 27, pp. 45-60, 2018.

[16] J. Tian, "Reversible data embedding using a difference expansion," *IEEE transactions on circuits and systems for video technology,* vol. 13, no. 8, pp. 890-896, 2003.

[17] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Transactions on circuits and systems for video technology,* vol. 16, no. 3, pp. 354-362, 2006.

[18] D. M. Thodi and J. J. Rodríguez, "Expansion embedding techniques for reversible watermarking," *IEEE transactions on image processing,* vol. 16, no. 3, pp. 721-730, 2007.

[19] Y. Luo, F. Peng, X. Li, and B. Yang, "Reversible image water marking based on prediction-error expansion and compensation," in *2011 IEEE International Conference on Multimedia and Expo*, 2011: IEEE, pp. 1-6.

[20] P. Pal, P. Chowdhuri, and T. Si, "A novel watermarking scheme for medical image using support vector machine and lifting wavelet transform," *Multimedia Tools and Applications,* vol. 82, no. 26, pp. 41187-41206, 2023.

[21] S. Rahman *et al.*, "A comprehensive study of digital image steganographic techniques," *IEEE Access,* vol. 11, pp. 6770-6791, 2023.

[22] C. Zhang, B. Ou, F. Peng, Y. Zhao, and K. Li, "A survey on reversible data hiding for uncompressed images," *ACM Computing Surveys,* vol. 56, no. 7, pp. 1-33, 2024.

[23] G. Coatrieux, L. Lecornu, B. Sankur, and C. Roux, "A review of image watermarking applications in healthcare," in *2006 International conference of the IEEE Engineering in Medicine and Biology Society*, 2006: IEEE, pp. 4691-4694.

[24] H. Nyeem, W. Boles, and C. Boyd, "A review of medical image watermarking requirements for teleradiology," *Journal of digital imaging,* vol. 26, pp. 326-343, 2013.

[25] S. M. Mousavi, A. Naghsh, and S. Abu-Bakar, "Watermarking techniques used in medical images: a survey," *Journal of digital imaging,* vol. 27, pp. 714-729, 2014.

[26] A. Odeh, A. A. Taleb, T. Alhajahjeh, and F. Navarro, "Invisible Shield: Unveiling an Efficient Watermarking Solution for Medical Imaging Security," *Applied Sciences,* vol. 13, no. 24, p. 13291, 2023.

[27] O. Evsutin, A. Melman, and R. Meshcheryakov, "Digital steganography and watermarking for digital images: A review of current research directions," *IEEE Access,* vol. 8, pp. 166589-166611, 2020.

[28] X. Xiong, "Novel scheme of reversible watermarking with a complementary embedding strategy," *IEEE Access,* vol. 7, pp. 136592-136603, 2019.

[29] A. Menendez-Ortiz, C. Feregrino-Uribe, R. Hasimoto-Beltran, and J. J. Garcia-Hernandez, "A survey on reversible watermarking for multimedia content: A robustness overview," *IEEE Access,* vol. 7, pp. 132662-132681, 2019.

[30] B. Ou, X. Li, Y. Zhao, R. Ni, and Y.-Q. Shi, "Pairwise prediction-error expansion for efficient reversible data hiding," *IEEE Transactions on image processing,* vol. 22, no. 12, pp. 5010-5021, 2013.

[31] X. Li, B. Yang, and T. Zeng, "Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection," *IEEE transactions on image processing,* vol. 20, no. 12, pp. 3524-3533, 2011.

[32] N. Mishra and S. Jain, "RHOMBUS PATTERN BASED REVERSIBLE WATERMARKING USING PAIRWISE PREDICTION ERROR EXPANSION," 2016.

[33] Z. Li *et al.*, "Nearly-lossless-to-lossy medical image compression by the optimized JPEGXT and JPEG algorithms through the anatomical regions of interest," *Biomedical Signal Processing and Control,* vol. 83, p. 104711, 2023.

[34] A. R. Khan, S. Khan, M. Harouni, R. Abbasi, S. Iqbal, and Z. Mehmood, "Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification," *Microscopy Research and Technique,* vol. 84, no. 7, pp. 1389-1399, 2021.

[35] H. S. Al-Dmour, *Enhancing information hiding and segmentation for medical images using novel steganography and clustering fusion techniques*. University of Technology Sydney (Australia), 2018.

[36] A. Kumar, "Study and analysis of different segmentation methods for brain tumor MRI application," *Multimedia Tools and Applications,* vol. 82, no. 5, pp. 7117-7139, 2023.

[37] R. E. Arevalo-Ancona and M. Cedillo-Hernandez, "Zero-watermarking for medical images based on regions of interest detection using K-means clustering and discrete fourier transform," *International Journal of Advanced Computer Science and Applications,* vol. 14, no. 6, 2023.

[38] M. I. Sarwar, L. A. Maghrabi, I. Khan, Q. H. Naith, and K. Nisar, "Blockchain: A crypto-intensive technology—A comprehensive review," *IEEE Access,* vol. 11, pp. 141926-141955, 2023.

[39] S. Dilhara, "A Review on Application of Hash Functions and Digital signatures in the Blockchain Industry," 2021.

[40] C. Gilbert and M. Gilbert, "Exploring Secure Hashing Algorithms for Data Integrity Verification," *Available at SSRN 5251606,* 2025.

[41] G. Mamatha and A. Shaik, "Reversible Data Hiding Based on Integer Wavelet Transforms and Prediction Error Expansion," in *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2018: IEEE, pp. 330-335.

[42] A. Shaik and V. Thanikaiselvan, "Comparative analysis of integer wavelet transforms in reversible data hiding using threshold based histogram modification," *Journal of King Saud University-Computer and Information Sciences,* vol. 33, no. 7, pp. 878-889, 2021.

[43] S. Yi, Y. Zhou, and Z. Hua, "Reversible data hiding in encrypted images using adaptive block-level prediction-error expansion," *Signal Processing: Image Communication,* vol. 64, pp. 78-88, 2018.

[44] N. O. Abokhdair *et al.*, "Patient data hiding and integrity control using prediction-based watermarking for brain MRI and CT scan images," *Journal of Medical Imaging and Health Informatics,* vol. 8, no. 4, pp. 691-702, 2018.

[45] A. M. Kapadia and P. Nithyanandam, "A Recursive High Payload Reversible Data Hiding Using Integer Wavelet and Arnold Transform," *Intelligent Automation & Soft Computing,* vol. 35, no. 1, 2023.

[46] S. Bhalerao, I. A. Ansari, and A. Kumar, "A reversible medical image watermarking for ROI tamper detection and recovery," *Circuits, Systems, and Signal Processing,* vol. 42, no. 11, pp. 6701-6725, 2023.

[47] A. Arham, H. A. Nugroho, and T. B. Adji, "Reversible Medical Image Watermarking for ROI-Based Tamper Detection and Recovery," 2023.

[48] H. Shi, Y. Wang, Y. Li, Y. Ren, and C. Guo, "Region-based reversible medical image watermarking algorithm for privacy protection and integrity authentication," *Multimedia Tools and Applications,* vol. 80, pp. 24631-24667, 2021.

[49] N. N. Hurrah, S. A. Parah, and J. A. Sheikh, "Embedding in medical images: an efficient scheme for authentication and tamper localization," *Multimedia Tools and Applications,* vol. 79, pp. 21441-21470, 2020.

[50] R. Geetha and S. Geetha, "Embedding electronic patient information in clinical images: an improved and efficient reversible data hiding technique," *Multimedia Tools and Applications,* vol. 79, no. 19, pp. 12869-12890, 2020.

[51] D. Li, X. Dai, J. Gui, J. Liu, and X. Jin, "A reversible watermarking for image content authentication based on wavelet transform," *Signal, Image and Video Processing,* vol. 18, no. 3, pp. 2799-2809, 2024.

[52] T.-S. Nguyen and P.-H. Vo, "Reversible image authentication scheme based on prediction error expansion," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 21, no. 1, pp. 253-262, 2021.

[53] S. Agrawal and M. Kumar, "Reversible data hiding for medical images using integer-to-integer wavelet transform," in *2016 IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, 2016: IEEE, pp. 1-5.

[54] C.-C. Lin, C.-C. Chang, W.-J. Kao, and J.-F. Chang, "Efficient electronic patient information hiding scheme with tamper detection function for medical images," *IEEE Access,* vol. 10, pp. 18470-18485, 2022.

[55] D. Awasthi and V. K. Srivastava, "Robust, imperceptible and optimized watermarking of DICOM image using Schur decomposition, LWT-DCT-SVD and its authentication using SURF," *Multimedia Tools and Applications,* vol. 82, no. 11, pp. 16555-16589, 2023.

[56] X. Liu *et al.*, "A novel robust reversible watermarking scheme for protecting authenticity and integrity of medical images," *Ieee Access,* vol. 7, pp. 76580-76598, 2019.

[57] R. Bamal and S. S. Kasana, "Reversible medical image watermarking for tamper detection using ANN and SLT," *Multimedia Tools and Applications,* vol. 83, no. 8, pp. 21849-21882, 2024.

[58] C. K. Tan, J. C. Ng, X. Xu, C. L. Poh, Y. L. Guan, and K. Sheah, "Security protection of DICOM medical images using dual-layer reversible watermarking with tamper detection capability," *Journal of Digital Imaging,* vol. 24, no. 3, pp. 528-540, 2011.

[59] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *Journal of Fourier analysis and applications,* vol. 4, no. 3, pp. 247-269, 1998.

[60] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 still image coding system: an overview," *IEEE transactions on consumer electronics,* vol. 46, no. 4, pp. 1103-1127, 2002.

[61] [Online]. Available: https://dicom.nema.org/medical/dicom/current/output/chtml/part05/sect_8.2.4.html.

[62] O. S. Pianykh, *Digital imaging and communications in medicine (DICOM) a practical introduction and survival guide*. Springer, 2012.

[63] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, 2003, vol. 2: Ieee, pp. 1398-1402.

[64] Y. Bai, G. Jiang, Z. Zhu, H. Xu, and Y. Song, "Reversible data hiding scheme for high dynamic range images based on multiple prediction error expansion," *Signal Processing: Image Communication,* vol. 91, p. 116084, 2021.

[65] A. F. Qasim, R. Aspin, F. Meziane, and P. Hogg, "ROI-based reversible watermarking scheme for ensuring the integrity and authenticity of DICOM MR images," *Multimedia Tools and Applications,* vol. 78, pp. 16433-16463, 2019.

[66] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing,* vol. 13, no. 4, pp. 600-612, 2004.

Appendices

Specifications of the medical images' dataset used

Image Name	Image Size	Modality	Bit depth
CTimg1	512 × 512	CT	12
CTimg2	512 × 512	CT	12
CTimg3	512 × 512	CT	12
CTimg4	512 × 512	CT	12
CTimg5	512 × 512	CT	12
CTimg6	512 × 512	CT	12
CTimg7	512 × 512	CT	12
CTimg8	512 × 512	CT	12
CTimg9	512 × 512	CT	12
CTimg10	512 × 512	CT	12
CTimg12	512 × 512	CT	12
CTimg16	512 × 512	CT	12
CTimg17	512 × 512	CT	12
CTimg18	512 × 512	CT	12
CTimg19	512 × 512	CT	12
CTimg20	512 × 512	CT	12
CTimg21	512 × 512	CT	12
CTimg22	512 × 512	CT	12
CTimg23	768 × 768	CT	12
CTimg24	512 × 512	CT	12
CTimg28	512 × 512	CT	12
CTimg29	512 × 512	CT	12
CTimg30	512 × 512	CT	12
CTimg33	512 × 512	CT	12

Image Name	Image Size	Modality	Bit depth
CTimg37	768 × 768	CT	12
CTimg38	512 × 512	CT	12
CTimg39	512 × 512	CT	12
CTimg40	512 × 512	CT	12
CTimg41	512 × 512	CT	12
CTimg42	512 × 512	CT	12
CTimg43	512 × 512	CT	12
CTimg44	512 × 512	CT	12
CTimg45	512 × 512	CT	12
MRIimg11	512 × 512	MR	12
MRIimg14	512 × 512	MR	12
MRIimg17	256 × 256	MR	12
MRIimg19	512 × 512	MR	12
MRIimg31	512 × 512	MR	12
MRIimg32	512 × 512	MR	12
MRIimg37	512 × 512	MR	16
MRIimg48	256 × 256	MR	12

Results of the Algorithm with Quality and Capacity Metrics

Image Name	ROI Size%	ROI Cap (bit)	ROI Cap (bpp)	PSNR (ROI)	RONI Cap (bit)	RONI Cap (bpp)	PSNR (RONI)	Total Cap (bit)	Cap (bpp)	Total PSNR	SSIM
CTimg1	43.6241	25267	0.1741	48.8294	172818	0.5349	66.5160	198085	0.7090	51.2616	0.9808
CTimg2	50.9861	16139	0.0870	42.3110	209644	0.7629	59.4958	225783	0.8499	45.3223	0.9672
CTimg3	44.2455	16934	0.0963	46.2471	219128	0.7133	63.1898	236062	0.8096	48.7151	0.9732
CTimg4	50.9861	16139	0.0870	42.3110	209644	0.7629	59.4958	225783	0.8499	45.3223	0.9672
CTimg5	50.3597	27231	0.1540	41.1691	204170	0.6987	60.6335	231401	0.8527	44.1612	0.9504
CTimg6	40.2046	28857	0.2125	48.3422	161688	0.4704	65.2661	190545	0.6829	50.5166	0.9832
CTimg7	50.1461	26028	0.1588	44.2523	189122	0.6355	64.8675	215150	0.7943	47.2375	0.9737
CTimg8	47.8352	26824	0.1689	48.2452	186944	0.6051	65.8065	213768	0.7740	51.0021	0.9854
CTimg9	49.3626	21369	0.1210	44.1333	196198	0.7067	61.9795	217567	0.8276	47.0196	0.9745
CTimg10	50.1461	26028	0.1588	44.2523	189122	0.6355	64.8675	215150	0.7943	47.2375	0.9737
CTimg12	47.8352	26824	0.1689	48.2452	186944	0.6051	65.8065	213768	0.7740	51.0021	0.9854
CTimg16	44.8196	17192	0.0989	46.2449	216486	0.7028	63.7097	233678	0.8018	48.7643	0.9725
CTimg17	30.2917	30592	0.1744	65.9852	246944	0.5359	70.0373	277536	0.5228	66.8249	0.9995
CTimg18	36.9068	38402	0.2049	64.3549	223116	0.6084	69.9468	261518	0.6218	65.6980	0.9995
CTimg19	45.3854	50744	0.2708	60.9122	182162	0.4963	69.7647	232906	0.8477	63.0552	0.9987
CTimg20	32.6084	12823	0.0684	52.6661	210014	0.5724	62.7700	222837	0.5675	54.1795	0.9912
CTimg21	26.9348	28600	0.3212	45.4745	100268	0.1594	51.1428	128868	0.4806	46.4236	0.9509
CTimg22	49.5697	20331	0.1154	43.9518	204164	0.7086	61.5008	224495	0.8240	46.8504	0.9741
CTimg23	43.9309	12395	0.0971	51.4182	21706	0.7997	53.7891	34101	0.8969	52.3056	0.9840
CTimg24	37.8380	22543	0.1727	52.2392	174516	0.4573	62.8098	197059	0.6300	54.0781	0.9929

Image Name	ROI Size%	ROI Cap (bit)	ROI Cap (bpp)	PSNR (ROI)	RONI Cap (bit)	RONI Cap (bpp)	PSNR (RONI)	Total Cap (bit)	Cap (bpp)	Total PSNR	SSIM
CTimg28	39.0106	17165	0.1285	52.2199	184680	0.5005	63.0672	201845	0.6290	54.1446	0.9913
CTimg29	47.8352	26824	0.1689	48.2452	186944	0.6051	65.8065	213768	0.7740	51.0021	0.9854
CTimg30	44.3703	16462	0.0965	46.2542	214768	0.6907	63.3632	231230	0.7873	48.7343	0.9734
CTimg33	26.9825	27727	0.3096	45.4558	102488	0.1712	51.1103	130215	0.4808	46.4056	0.9509
CTimg38	50.1133	26947	0.1628	44.2270	193574	0.6416	63.3501	220521	0.8045	47.1941	0.9729
CTimg39	50.9964	18271	0.0998	40.9511	208088	0.7486	57.8543	226359	0.8484	43.9575	0.9567
CTimg40	49.3626	21369	0.1210	44.1333	196198	0.7067	61.9795	217567	0.8276	47.0196	0.9745
CTimg41	50.3597	27231	0.1540	41.1691	204170	0.6987	60.6335	231401	0.8527	44.1612	0.9504
CTimg42	49.7704	26872	0.1558	44.3020	198240	0.6789	62.4588	225112	0.8347	47.2271	0.9708
CTimg43	44.1921	24210	0.1598	46.3053	210628	0.5727	65.4466	234838	0.7325	48.7966	0.9722
CTimg44	48.4505	16673	0.0913	44.3661	207738	0.7472	62.6809	224411	0.8385	47.1841	0.9697
MRIimg11	37.2757	14444	0.1131	40.5059	207532	0.7334	46.3736	221976	0.8465	41.9099	0.7768
MRIimg14	33.0853	16967	0.1504	42.5934	192612	0.6142	53.1023	209579	0.7646	44.1512	0.8205
MRIimg17	31.9336	11450	0.4210	45.2628	45674	0.4296	46.7208	57124	0.8506	45.6775	0.8580
MRIimg19	31.4880	17663	0.1613	42.6129	198458	0.5871	52.9513	216121	0.7484	44.0744	0.8179
MRIimg31	28.3409	16553	0.1790	45.4103	153782	0.4569	51.0920	170335	0.6359	46.4165	0.8797
MRIimg32	33.0853	16967	0.1504	42.5934	192612	0.6142	53.1023	209579	0.7646	44.1512	0.8205
MRIimg37	39.6538	13153	0.0972	58.6616	232094	0.7998	68.9048	245247	0.8970	60.5933	0.9925
MRIimg48	37.5610	11862	0.3797	45.1642	50958	0.5768	47.1725	62820	0.9565	45.8145	0.8744

Extraction Results with Tamper Detection

Image Name	Extracted EPR (%)	Tampered	Authentic ROI	Authentic RONI	PSNR (dB)	SSIM
CTimg1	100	No	yes	yes	inf	1.0000
CTimg2	100	No	yes	yes	inf	1.0000
CTimg3	100	No	yes	yes	inf	1.0000
CTimg4	100	No	yes	yes	inf	1.0000
CTimg5	100	No	yes	yes	inf	1.0000
CTimg6	100	No	yes	yes	inf	1.0000
CTimg7	100	No	yes	yes	inf	1.0000
CTimg8	100	No	yes	yes	inf	1.0000
CTimg9	100	No	yes	yes	inf	1.0000
CTimg10	100	No	yes	yes	inf	1.0000
CTimg12	100	No	yes	yes	inf	1.0000
CTimg16	100	No	yes	yes	inf	1.0000
CTimg22	100	No	yes	yes	inf	1.0000
CTimg24	100	No	yes	yes	inf	1.0000
CTimg28	100	No	yes	yes	inf	1.0000
CTimg29	100	No	yes	yes	inf	1.0000
CTimg30	100	No	yes	yes	inf	1.0000
CTimg33	100	No	yes	yes	inf	1.0000
CTimg37	100	Yes	No	yes	90.06	0.999999
CTimg37	100	Yes	No	yes	89.13	0.999999
CTimg38	100	No	yes	yes	inf	1.0000
CTimg39	100	Yes	No	yes	83.89	0.999970
CTimg40	100	No	yes	yes	inf	1.0000
CTimg41	100	No	yes	yes	inf	1.0000
CTimg42	100	No	yes	yes	inf	1.0000
CTimg43	100	Yes	No	yes	87.57	0.999996
CTimg44	100	Yes	No	yes	129.98	1.0000

Image Name	Extracted EPR (%)	Tampered	Authentic ROI	Authentic RONI	PSNR (dB)	SSIM
CTimg45	100	Yes	No	yes	89.16	0.999993
MRIimg11	100	No	yes	yes	inf	1.0000
MRIimg11	100	Yes	No	yes	73.09	0.999983
MRIimg11	97.49	Yes	No	yes	98.63	0.999999
MRIimg14	100	No	yes	yes	inf	1.0000
MRIimg17	100	No	yes	yes	inf	1.0000
MRIimg17	69.93	Yes	No	yes	102.87	0.999999
MRIimg17	89.66	Yes	No	yes	68.73	0.999795
MRIimg19	100	No	yes	yes	inf	1.0000
MRIimg31	100	No	yes	yes	inf	1.0000
MRIimg32	100	No	yes	yes	inf	1.0000
CTimg21	100	No	yes	yes	inf	1.0000
MRIimg37	100	Yes	No	yes	76.71	0.999995
CTimg20	100	Yes	No	yes	97.13	0.999999
CTimg1	100	Yes	No	yes	94.80	0.999999
MRIimg17	93.34	Yes	No	yes	102.17	0.999999

المستخلص

أدى الانتشار الواسع للتطبيب عن بُعد وتبادل الصور الطبية عبر الشبكات العامة إلى مخاوف بشأن أمن وسلامة بيانات المرضى الحساسة. وقد برزت تقنية العلامات المائية الرقمية كتقنية واعدة لتأمين الصور الطبية المتبادلة وإخفاء معلومات المرضى، مما يوفر السرية والتحقق من الهوية وكشف التلاعب. وتُعدّ طرق العلامات المائية القابلة للعكس ذات قيمة خاصة للتطبيقات الطبية، إذ تُمكّن من استعادة الصورة الأصلية بالكامل بعد استخراج العلامة المائية، مما يحافظ على جودة التشخيص. ولتعزيز المتانة ضد الهجمات، وتحقيق سعة عالية، وضمان عدم إمكانية اكتشافها، يُقدّم هذا البحث مخططًا جديدًا للعلامات المائية القابلة للعكس لصور الرنين المغناطيسي والتصوير المقطعي المحوسب للدماغ. ويُدمج هذا المخطط تحويل المويجات الصحيحة القائم على الرفع مع توسيع خطأ التنبؤ التكيفي لضمان سلامة الصور الطبية ومعلومات المرضى ومصداقيتها وسريتها. ولتحسين كفاءة التضمين، يُقسّم المخطط معاملات الصورة الطبية باستخدام تقنية التعلم الآلي غير الخاضعة للإشراف إلى منطقة الاهتمام (ROI) ومنطقة عدم الاهتمام (RONI). للحفاظ على جودة المنطقة ذات الأهمية التشخيصية، يتم إخفاء بت واحد في كل معامل قابل للتضمين في منطقة الاهتمام (ROI)، بينما يتم إخفاء بتّين في كل معامل قابل للتضمين في منطقة عدم الاهتمام (RONI). يتم تضمين قيمة التجزئة لمنطقة الاهتمام وبيانات المريض في منطقة الاهتمام، بينما يتم تضمين آلية كشف التلاعب واستعادته وقيمة التجزئة لمنطقة عدم الاهتمام في هذه المنطقة. تم تقييم متانة النظام المقترح في مواجهة هجمات الضوضاء الغاوسية بمستويات شدة مختلفة والهجمات المتعمدة، بما في ذلك القطع والقص والنسخ واللصق. تُؤكد النتائج التجريبية كفاءة النظام المقترح، وقابليته للعكس، والتحقق من سلامة البيانات، وحماية الخصوصية والأمان، وسعة التضمين العالية مع جودة صورة عالية. بالإضافة إلى ذلك، يكشف النظام بدقة المناطق المُتلاعب بها ويستعيدها بجودة عالية.