Democratic and Popular Republic of Algeria

Ministry of Higher Education and Scientific Research

University of Kasdi Merbah, Ouargla

Faculty of new information and communication technologies

Department of computer science and Information technology

Thesis Submitted to the Department of Computer Science and Information Technology in Candidacy for the Degree of "Doctor" 3rd Cycle LMD in Computer Science.

Deep versus handcrafted approaches for improving medical image security using watermarking

Presented by:

Khaled HEBBACHE

Committee members

President	Akram Zineddine Boukhamla	MCA- Ouargla University
Supervisor	Oussama Aiadi	MCA- Ouargla University
Co-Supervisor	Belal Khaldi	MCA- Ouargla University
Examiner	Bachir Said	MCA- Ouargla University
Examiner	Hocine Belouaar	MCA- Biskra University
Examiner	Abdehakim Cheriet	MCA- The National School of Artificial Intelligence (ENSIA) -Alger

Academic Year: 2024/2025

Dedication

To my dearest parents, brothers, and sisters

To my beloved wife, Meriem

To my adorable children

Abdelerrahmane

and Oussama

to all my family, To all my friends

To all those who were giving me any kind of support.

Acknowledgments

First and foremost, I would like to express my profound thanks to God for granting me the strength, courage, and health to complete this work.

I extend my heartfelt thanks to my supervisors, Dr. Oussama Aiadi and Dr. Belal Khaldi, for their invaluable guidance, constant support, and wise counsel throughout this journey. Their expertise and encouragement have been instrumental in shaping this research.

I am sincerely grateful to the jury members for their time and consideration in evaluating my work. Special thanks to Dr. Abdehakim Cheriet, Dr. Hocine Boulouaar, and Dr. Bachir Said for serving as examiners, and to Dr. Akram Zineddine Boukhamla for presiding as the head of the jury.

I would also like to acknowledge the professors in the Department of Computer Science and Information Technology, and all those whose knowledge and mentorship have been essential throughout my academic journey. I am deeply appreciative of all the members of our research group and my colleagues for their collaboration and support.

Finally, I express my gratitude to everyone who, in one way or another, contributed to the successful completion of this work. Your support has been invaluable.

Abstract

As advancements in computer vision continue to transform healthcare, the need to protect medical images, particularly in telemedicine, has become increasingly critical. This thesis explores the domain of medical image watermarking, offering a detailed comparison between handcrafted methods and deep learning approaches, with a specific emphasis on feature extraction techniques. A comprehensive state-of-the-art review sets the stage, identifying the strengths and limitations of various watermarking strategies aimed at safeguarding medical images.

The primary contributions of this research include three proposed watermarking methods. First, a novel blind watermarking method based on Local Binary Patterns and Discrete Wavelet Transform (LBP-DWT) is presented, specifically designed for telemedicine applications. Second, a gradient-based feature extraction technique, termed "GradWater," is developed to further strengthen watermark security through rich image-driven features. Third, a deep learning-based zero watermarking technique is proposed, utilizing a pre-trained VGG16 model. This method generates the watermark without embedding any alterations into the original medical image, ensuring its quality remains intact while maintaining strong security.

An extensive experimental evaluation compares the feature extraction capabilities of handcrafted methods against those of deep learning approaches, focusing on their resistance to various attacks, such as noise, compression, and geometric distortions, while preserving the quality of the medical images. The findings also offer a comparative analysis of zero and non-zero watermarking schemes, providing valuable insights into their respective advantages.

Keywords : Medical Image Watermarking Handcrafted Methods, Deep Learning, Local Binary Patterns, Discrete Wavelet Transform, Gradient, VGG16, Feature Extraction.

Résumé

Avec les avancées de la vision par ordinateur qui continuent de transformer le domaine de la santé, la nécessité de protéger les images médicales, en particulier dans la télémédecine, devient de plus en plus cruciale. Cette thèse explore le domaine du tatouage numérique des images médicales, en offrant une comparaison détaillée entre les méthodes artisanales et les approches basées sur l'apprentissage profond, avec un accent particulier sur les techniques d'extraction de caractéristiques. Une revue complète de l'état de l'art est présentée, identifiant les forces et les limites des différentes stratégies de tatouage visant à protéger les images médicales.

Les principales contributions de cette recherche comprennent trois méthodes de tatouage numérique proposées. Premièrement, une nouvelle méthode de tatouage aveugle basée sur les motifs binaires locaux (LBP) et la transformation en ondelettes discrète (DWT) est présentée, spécifiquement conçue pour les applications de télémédecine. Deuxièmement, une technique d'extraction de caractéristiques basée sur le gradient, appelée "GradWater", est développée pour renforcer davantage la sécurité des tatouages à travers des caractéristiques riches et tirées des images. Troisièmement, une technique de tatouage zéro basée sur l'apprentissage profond est proposée, utilisant un modèle pré-entraîné VGG16. Cette méthode génère le tatouage sans introduire de modifications dans l'image médicale originale, assurant ainsi la préservation de sa qualité tout en maintenant une sécurité robuste.

Une évaluation expérimentale approfondie compare les capacités d'extraction de caractéristiques des méthodes artisanales à celles des approches d'apprentissage profond, en se concentrant sur leur résistance à diverses attaques, telles que le bruit, la compression et les distorsions géométriques, tout en préservant la qualité des images médicales. Les résultats offrent également une analyse comparative des schémas de tatouage zéro et non-zéro, fournissant des informations précieuses sur leurs avantages respectifs.

Mots-clés : Tatouage d'images médicales, Méthodes artisanales, Apprentissage profond, Motifs Binaires Locaux, Transformation en ondelettes discrète, Gradient, VGG16, Extraction de caractéristiques.

الملخص

مع استمرار التقدم في رؤية الحاسوب في تطوير مجال الرعاية الصحية، أصبحت الحاجة إلى حماية الصور الطبية، لا سيما في مجال التطبيب عن بعد، أكثر إلحاحًا. تستكشف هذه الأطروحة مجال إخفاء العلامات المائية في الصور الطبية، حيث تقدم مقارنة مفصلة بين الطرق اليدوية والأساليب المعتمدة على التعلم العميق، مع التركيز بشكل خاص على تقنيات استخراج السمات. تم تقديم مراجعة شاملة لأحدث ما توصل إليه العلم، والتي تحدد نقاط القوة والقيود في استراتيجيات العلامات المائية المختلفة التي تهدف إلى حماية الصور الطبية.

تشمل المساهمات الرئيسية لهذا البحث ثلاثة أساليب مقترحة لإخفاء العلامات المائية. أولاً، يتم تقديم طريقة جديدة لإخفاء العلامات المائية العمياء تعتمد على الأنماط الثنائية المحلية (LBP) وتحويل المويجات المتقطعة(DWT) ، المصممة خصيصًا لتطبيقات التطبيب عن بعد. ثانيًا، تم تطوير تقنية استخراج السمات القائمة على التدرج، المسماة "GradWater"، لتعزيز أمان العلامة المائية من خلال السمات المستخرجة من الصور. ثالثًا، تم اقتراح تقنية العلامة المائية الصفرية المعتمدة على التعلم العميق باستخدام نموذج VGG16 المدرب مسبقًا. تولد هذه الطريقة العلامة المائية دون إدخال أي تعديلات على الصورة الطبية الأصلية، مما يضمن الحفاظ على جودتها مع الحفاظ على مستوى أمان قوي.

يتم إجراء تقييم تجريبي شامل لمقارنة قدرات استخراج السمات بين الأساليب اليدوية وتلك القائمة على التعلم العميق، مع التركيز على مقاومتها لهجمات مختلفة مثل الضوضاء، الضغط، والتشوهات الهندسية، مع الحفاظ على جودة الصور الطبية. كما تقدم النتائج تحليلًا مقارنًا بين أساليب العلامة المائية الصفرية وغير الصفرية، مما يوفر رؤى قيمة حول المزايا النسبية لكل منها.

الكلمات المفتاحية **:** إخفاء العلامات المائية في الصور الطبية، الأساليب اليدوية، التعلم العميق، الأنماط الثنائية المحلية، تحويل المويجات المتقطعة، التدرج، VGG16، استخراج السمات.

Table of Contents

List of Figures

List of Tables

I. General Introduction

I.1. **Introduction**

Computer vision, a burgeoning field within artificial intelligence, empowers machines to interpret and comprehend the visual world. It equips computers with the ability to extract meaningful information from images and videos, such as recognizing objects, tracking motion, and understanding scenes. This technology has found widespread applications across diverse domains, including medical imaging, autonomous vehicles, surveillance systems, and more [1]. One of the pivotal applications of computer vision is image watermarking. Watermarking involves discreetly embedding invisible data (watermarks) into images to safeguard intellectual property, verify authenticity, and track usage. By leveraging computer vision techniques, we can develop robust and imperceptible watermarking algorithms that can withstand various attacks and maintain the integrity of digital content[2].

Image protection, facilitated by techniques like watermarking, offers a plethora of benefits, including:

- Intellectual property preservation: Watermarking acts as a deterrent against unauthorized copying, distribution, and modification of images, safeguarding the rights of creators and owners.
- Authenticity verification: Watermarks serve as a reliable means to verify the authenticity of images, ensuring that they have not been tampered with or altered.
- Usage tracking: Watermarks enable the tracking of image distribution and usage, providing valuable insights for licensing, marketing, and analytics purposes.
- Forensic analysis: In cases of copyright infringement or image misuse, watermarks can aid in forensic investigations by identifying the source of the image.

Watermarking stands out as a particularly effective technique for image protection due to its several advantages, Watermarks can be embedded into images in a manner that is imperceptible to the human eye, ensuring that the visual quality of the image remains pristine. Watermarking techniques can be engineered to be resistant to various attacks, such as compression, noise, and geometric transformations, making it challenging for malicious actors to remove or alter the watermark [3]. Watermarking can be applied to a broad spectrum of image formats and types, making it a versatile tool for protecting digital content. Watermarking techniques can be easily adapted to accommodate large volumes of images, making them suitable for both individual and enterprise-level applications.

In the medical field, securing medical images is paramount for several reasons, medical images contain sensitive patient information, including personal identifiers and medical history. Protecting these images helps maintain patient privacy and prevent unauthorized disclosure of sensitive data. ampered or corrupted medical images can lead to inaccurate diagnoses and treatment decisions, potentially harming patient health[3]. Healthcare organizations are subject to various legal and regulatory requirements related to data security and privacy, watermarking can help ensure compliance with these regulations. Medical images are often used for research purposes. Protecting the integrity of these images is essential for ensuring the reliability and validity of research findings.

The COVID-19 pandemic has further accelerated the adoption of telemedicine, prompting healthcare providers to rely heavily on digital platforms for remote diagnosis and treatment. This rapid digital transformation has underscored the importance of securing medical data, particularly medical images, which are often transmitted across networks [4]. In this context, watermarking provides a valuable solution for verifying image authenticity, detecting tampering, and maintaining the confidentiality of sensitive data. Watermarking offers a comprehensive and effective solution for protecting medical images. Its ability to preserve diagnostic information, enhance security and authenticity, and integrate seamlessly with existing medical imaging workflows makes it a valuable tool for safeguarding sensitive patient data and ensuring the integrity of medical imaging systems [5].

Previous studies on image watermarking have focused on various aspects, including:

- Embedding mechanisms**:** Some approaches have focused on improving the mechanisms used to embed watermarks into images, such as developing more robust and imperceptible embedding techniques.
- Optimal region selection**:** Others have focused on identifying the optimal regions within an image for embedding watermarks, ensuring that the watermark is not easily detectable or removable.
- Watermark extraction**:** Some studies have investigated techniques for extracting watermarks from images, ensuring that the watermark can be reliably recovered even in the presence of noise or attacks.
- Feature extraction**:** A key focus area in our research is feature extraction, which involves extracting relevant features from images that can be used for watermarking and authentication.

Recent years have witnessed the emergence of a new category of watermarking techniques known as zero-watermarking. Unlike traditional watermarking schemes that embed watermarks directly into the image data, zero-watermarking techniques rely on hashing algorithms to generate a unique digital fingerprint for each image, enabling authentication and verification without modifying the original image content [6].

Investigating the feature extraction stage is essential in watermarking as it significantly impacts both imperceptibility and robustness. Techniques like Local Binary Patterns (LBP) and gradient analysis focus on extracting distinct image features such as textures and edges, providing optimal locations for embedding the watermark without degrading image quality. In deep learning-based watermarking, learned features from convolutional layers capture more complex and abstract image characteristics, enhancing the system's ability to embed watermarks robustly while maintaining visual fidelity. Effective feature extraction, therefore, ensures better resistance to attacks and higher accuracy in watermark retrieval [7].

Handcrafted watermarking techniques have undergone significant advancements, with two primary categories emerging: spatial domain techniques, which operate directly on image pixels (e.g., least significant bit (LSB) embedding), and transform domain techniques, which embed watermarks in transformed representations (e.g., discrete cosine transform (DCT), discrete wavelet transform (DWT)).

The field of image watermarking has recently witnessed a shift from handcrafted approaches to deep learning-based techniques (e.g., Convolutional Ceural Cetworks (CNNs)). While traditional handcrafted methods relied on manually designed algorithms, deep learning offers a data-driven approach that can automatically learn complex patterns and features from large datasets. While both handcrafted and deep learning-based approaches have made significant contributions to the field of watermarking, deep learning has gained significant traction in recent years due to its ability to automatically learn complex patterns and features from data. This has led to the development of more efficient watermarking algorithms [8], [9].

I.2. **Research Questions**

Feature extraction plays a pivotal role in watermarking systems, particularly for medical images where both data security and image quality are paramount. This work focuses

on understanding how different methods, both traditional and deep learning-based, utilize feature extraction to improve the robustness and imperceptibility of watermarks. By addressing the following research questions, we aim to explore how extracting key image features can lead to more secure and efficient watermarking techniques. In the current work, we can summarize the research questions as follows:

1/ Investigate the effect of feature extraction on the performance of watermarking schemes by studying the performance of several feature schemes. *How does the choice of feature extraction method impact the effectiveness of a watermarking system?* This question aims to explore the critical role feature extraction plays in watermark embedding and extraction processes, especially in the context of robustness against common attacks such as noise, compression, or geometric attacks.

2/ Studying the performance of features with more abundant information compared to the straightforward one. *How do more complex feature extraction methods, such as deep learning features, compare to simpler methods like gradient-based and LBP techniques in improving watermarking robustness?* This question is particularly focused on the contrast between low-level features (such as those captured by LBP and image gradients) and higher-level features (such as deep features).

3/ Investigate the performance of deep-learning-based features (CNN in specific) compared to the handcrafted features . *How do deep learning features, extracted from convolutional neural networks (CNNs), enhance watermarking performance compared to traditional handcrafted features?* In this question, the goal is to evaluate whether the rich, hierarchical features learned by CNNs can outperform handcrafted methods like LBP, DWT, or gradient-based approaches.

4/ Compare the performance of zero and non-zero watermarking schemes . *What are the key differences in performance between zero and non-zero watermarking schemes, and how do they impact security and robustness?* This question delves into two fundamental categories of watermarking schemes: zero-watermarking, where no alteration to the host image occurs, and non-zero watermarking, where the image is modified to embed the watermark.

5/ What are the limitations of the current watermarking techniques in medical imaging, and the future directions we can follow to address these limitations? This question aims to critically assess the current state of medical image watermarking and identify the limitations

that remain unsolved. For instance, existing techniques may face challenges with balancing robustness and imperceptibility, ensuring security in the presence of sophisticated attacks, or handling large-scale medical datasets with varying characteristics. Additionally, the question will explore potential future research directions, such as integrating advanced machine learning models, enhancing watermark security through novel feature extraction techniques, or developing hybrid methods that combine both traditional and deep learning-based approaches. This analysis will help guide future innovations in the field, focusing on improving the adaptability, robustness, and usability of watermarking systems in real-world medical applications.

I.3. **Overview of Related Works**

Numerous studies have been conducted to address the research questions outlined in this work, exploring various watermarking techniques in the context of medical imaging. The literature can be categorized into three main approaches: handcrafted methods, deep learning-based methods, and hybrid techniques that combine elements of both. Each category presents distinct methodologies, advantages, and challenges, particularly regarding the use of feature extraction methods in handcrafted approaches.

1.2.1. **Handcrafted Approaches**

Handcrafted approaches often leverage feature extraction methods to enhance watermarking techniques. Many of these methods either directly used the extracted features to protect the watermark in the spatial domain of the original image, resulting in fragile watermarking, or in the frequency domain, which yields robust watermarking[10] . For instance, fragile watermarking techniques typically use methods such as LBP, Zernike moments , Vector Map to ensure that any modification to the image will result in a noticeable loss of watermark integrity, making them ideal for tamper detection ([11], [12], [13]). Conversely, robust watermarking techniques may employ Scale-Invariant Feature Transform (SIFT), DWT or LBP to embed watermarks into the frequency domain, allowing for resilience against common image attacks ([14], [15]). Alternatively, some approaches extract features and merge them with the watermark(e,g,. using XOR) without embedding them into the image, leading to a zero watermarking scenario, which retains the original image's quality while still providing a level of protection ([16], [17], [18], [19]).

1.2.2. Deep Learning-Based Approaches

Deep learning approaches often utilize feature extraction within neural networks to enhance watermarking techniques. Typically, these methods automatically learn high-level features using convolutional neural networks (CNNs) [9], which can then be used to modify the image's feature space. This allows for robust watermarking, where the watermark is embedded into the feature domain rather than the pixel domain, providing significant resistance to attacks such as noise or compression([20], [21]). In some approaches, pre-trained deep learning models are employed for zero-watermarking, where the watermark key and the features learned by the network are used to generate a master share without modifying the original image, ensuring the image quality is preserved while maintaining security ([9], [22], [23]). These approaches are particularly suited to complex scenarios such as medical image protection and authentication.

1.2.3. Hybrid Approaches

Hybrid methods combine traditional handcrafted techniques with deep learning to enhance the performance and robustness of watermarking. These methods typically employ traditional transformations such as DWT, Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), or Singular Value Decomposition (SVD), followed by deep learning models such as Convolutional Neural Networks (CNNs) to refine the watermark embedding process, or vice versa([8], [9]). The watermark is then embedded either in the frequency domain or within the extracted features. This combination offers the advantages of both approaches: the precision and stability of traditional transformations and the flexibility of deep learning. In robust hybrid watermarking, the watermark is embedded into the feature domain using both handcrafted and learned features, providing enhanced resistance to distortions such as compression, noise, or geometric transformations([24], [25]). Additionally, some hybrid methods leverage pre-trained deep learning networks for zero-watermarking, where features extracted by traditional methods are combined with learned features to create a secure master share, all while leaving the host image unaltered([26], [27]).

It is worth noting that in this section we give only an overview on the related work, while we devote a whole chapter (Chapter 2) to detail the different aspects of methods concerned with watermarking.

I.4. **Motivations**

This thesis is motivated by the critical need to enhance the protection and authentication of medical images, particularly in telemedicine applications where data integrity is paramount. The increasing complexity of medical data in digital environments calls for more robust watermarking solutions to secure patient information. Exploring the rich information available in images allows for deriving more reliable features that can enhance security. While traditional handcrafted watermarking methods have demonstrated their effectiveness in balancing imperceptibility and robustness, there is growing interest in deep learning approaches due to their potential for adaptive feature extraction and improved security against sophisticated attacks.

An essential part of this research is the empirical comparison between handcrafted and deep learning-based watermarking techniques under the same experimental protocol. This investigation aims to determine whether learned features from deep models, such as pre-trained CNNs, can capture and reflect image content more effectively than handcrafted methods like LBP or gradient analysis. By examining these methods in the context of medical image processing, this work also aims to offer new insights into the role of feature extraction for watermark embedding and extraction.

I.5. **Contributions**

This research makes significant contributions to the field of medical image watermarking, addressing both theoretical and practical aspects of watermarking techniques. The following key contributions are presented:

1. A comprehensive state-of-the-art review that evaluates the advantages and disadvantages of various watermarking approaches. This review provides a critical analysis of existing methods, highlighting their effectiveness in terms of robustness, imperceptibility, and computational efficiency. By synthesizing recent developments in both handcrafted and deep learning-based techniques, this contribution aims to establish a clear understanding of the current landscape in medical image watermarking.
2. The proposal of a new handcrafted approach for medical image watermarking based on efficient Local Binary Patterns (LBP) features. This method leverages the spatial

structure of LBP to enhance watermark embedding while maintaining image quality. By optimizing the feature extraction process, this contribution aims to improve robustness against common attacks and increase the reliability of watermark detection in medical applications.

3. Introduction of a novel gradient-based feature for image watermarking, termed "GradWater." This feature utilizes gradient analysis to capture intricate details within medical images, thereby enhancing watermark security. By generating a rich set of image-driven features, GradWater can adaptively adjust to various image contents, leading to improved robustness and imperceptibility of the watermarked images.
4. An experimental investigation comparing the performance of handcrafted versus deep learning-based approaches for medical image watermarking. This study rigorously evaluates the effectiveness of both methodologies under controlled conditions, aiming to determine the strengths and limitations of each approach. Insights gained from these experiments will inform best practices for implementing watermarking systems in medical settings.
5. An experimental analysis of the performance of zero and non-zero watermarking schemes. This investigation explores the trade-offs associated with each method, assessing factors such as extraction accuracy, imperceptibility, and robustness against attacks. The findings will contribute to a deeper understanding of how different watermarking strategies can be applied effectively in medical imaging scenarios.

I.6. **Thesis Structure**

This thesis is organized into five chapters:

- Chapter 1 introduces the fundamental principles of digital images and basic concepts of image watermarking. It covers key image properties such as formats (e.g., JPG, PNG), resolution, and noise, and discusses essential aspects of watermarking, including imperceptibility, robustness, and capacity. This chapter establishes the foundational knowledge necessary for understanding subsequent technical discussions.
- Chapter 2 provides a comprehensive review of state-of-the-art approaches in image watermarking. It explores traditional handcrafted techniques, such as DWT and LBP, and compares them with recent deep learning-based methods, including CNNs. The

comparison emphasizes their strengths and limitations in terms of robustness, security, and computational efficiency.

- Chapter 3 introduces the proposed methods, representing the core contributions of the thesis. The first contribution is an LBP-based watermarking scheme tailored for medical images. The second contribution is a gradient-based watermarking approach designed to improve robustness against various attacks. Finally, the chapter presents a deep learning-based Zero-watermarking technique leveraging CNNs for effective watermark embedding and extraction. Each method is discussed in detail, highlighting its innovations and benefits.
- Chapter 4 focuses on experimental results and analysis. It outlines the experimental setup, including datasets and evaluation metrics, such as Peak Signal-to-Noise Ratio (PSNR) and Normalized Correlation (NC). The chapter evaluates the performance of the proposed methods under different attack scenarios, such as noise, blurring, and resizing, and offers a detailed comparison of the results. The findings are critically analyzed to assess the strengths and limitations of each approach.
- Chapter 5 concludes the thesis by summarizing key contributions and results. It reflects on the limitations of the current work and suggests avenues for future research, particularly in exploring further deep learning techniques and their potential applications in real-time medical image security.

At the end of the thesis, we draw the main conclusions of the work and we introduce some perspectives and future works.

II. Chapter 1 : Image Watermarking: Basic Principles and Concepts

II.1. **Introduction:**

In this chapter, we delve into the foundational principles of image watermarking, a critical tool for ensuring the security and integrity of digital images, particularly in medical contexts. The chapter begins by outlining the basic characteristics of digital images, including their types, formats, and importance in various applications, before transitioning into the core concepts of watermarking. The fundamental processes involved in watermark embedding and extraction are explored, alongside different types of watermarking techniques. A thorough discussion on key properties such as robustness, imperceptibility, and the impact of various watermarking attacks will provide a solid grounding for understanding how watermarking can be applied effectively in different scenarios. This framework sets the stage for the more advanced discussions and experimental evaluations in later chapters.

II.2. **Basic of Digital Images:**

II.2.1. Definition of Image

In computer science, an image is a digital representation of visual data in a two-dimensional array of pixels. Each pixel contains information about color or intensity, forming a grid of points that represent an object, scene, or pattern (Fig. 1) [6]. These images can be stored, manipulated, and processed by digital systems. The fundamental concept of an image relies on capturing light (or another type of wave) and converting it into a measurable quantity that can be stored and interpreted by computers [28].

Fig. 1 represents a matrix of numerical values for a gray image.

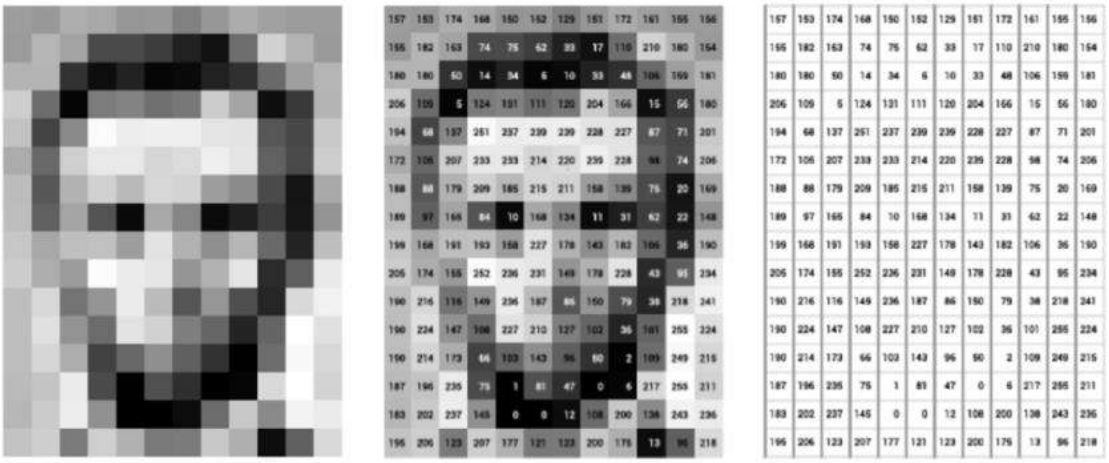


Fig 1.Digital image as matrix of numerical values[29].

Images play a crucial role in various domains, including medical diagnostics, artificial intelligence, entertainment, and security [3]. For digital systems, the understanding and processing of these images are critical for interpreting visual data accurately.

II.2.2. Types of Images

Images can be classified based on their content and color representation (Fig. 2), with each type serving different applications and processing requirements:

- **Grayscale Images**: These are images where each pixel represents a shade of gray. Typically, the pixel values range from 0 (black) to 255 (white), with various shades of gray in between. Grayscale images are widely used in image processing tasks because they focus on intensity values without the complexity of color information [30].

- **Color Images (RGB)**: RGB images consist of three color channels Red, Green, and Blue. Each pixel is represented by a combination of these three colors, allowing for the full spectrum of visible colors to be displayed. RGB images are standard in applications requiring detailed color representation, such as photography and video[6].

- **Binary Images**: A binary image consists of only two possible pixel values: 0 (black) and 1 (white). These images are commonly used in tasks that require simple on/off or presence/absence information, such as text recognition or edge detection [31].

- **Indexed Color Images**: In indexed color images, pixel values represent indices into a fixed color palette, reducing the amount of data required to store the image. This format is typically used in applications that need to compress image data without significant loss of visual information [32].

- **Medical Images**: Medical images are digital representations of the internal structures of a patient's body. Common modalities include X-rays, CT scans, MRI, ultrasound, and more. These images are essential for diagnostic purposes and are typically grayscale to highlight subtle tissue differences. Medical images often have high resolutions to allow doctors to make accurate assessments of conditions [33].

Fig 2. Basic types of digital images

II.2.3. Importance of Images

Images are essential in modern computing and digital technologies. Their significance extends to fields such as:

- **Image Processing**: Enhancing, filtering, and transforming images to improve their quality or extract specific features. This is fundamental for applications such as satellite imaging, photography, and medical diagnostics [34].

- **Computer Vision**: Enabling computers to understand and interpret visual data. Image recognition, object detection, and facial recognition rely on this field [32].

- **Medical Imaging**: Medical images are indispensable for healthcare. They allow for non-invasive examination of internal structures, helping in early disease detection, surgical planning, and monitoring of ongoing treatment. The precision of medical images can be life-saving as it enables clinicians to detect anomalies that are not visible in physical exams [35].

- **Security**: Images are used in biometric systems, surveillance, and watermarking to protect digital content and ensure secure identification processes [36].

The importance of images, especially medical images, cannot be overstated. They are not only tools for visualizing information but also central to decision-making in fields like

healthcare and security [36]. Advanced processing techniques ensure that the information extracted from these images is as accurate and reliable as possible [37].

II.2.4. Image Formats

The format of an image affects how it is stored, processed, and transmitted. Different image formats are suited for various applications based on factors like quality, compression, and compatibility:

- **JPEG (Joint Photographic Experts Group)**: This is a commonly used format for compressing photographic images. It uses lossy compression, which reduces file size at the expense of some image quality. JPEG is widely used for web images and photography but may not be ideal for tasks requiring lossless data retention[6].

- **PNG (Portable Network Graphics)**: is a lossless compression format, meaning that the image quality is preserved without any loss of information. PNG supports transparency and is often used in applications where image quality is critical, such as in medical imaging or digital art.

- **TIFF (Tagged Image File Format)**: is a flexible format that supports both lossless and lossy compression. It is frequently used in professional photography and medical imaging due to its ability to store high-quality images with detailed metadata.

- **BMP (Bitmap Image File)**: A simple raster image format that stores pixel data without compression. BMP files are large but preserve image quality, making them suitable for high-fidelity applications where compression artifacts are undesirable.

- **DICOM (Digital Imaging and Communications in Medicine)**: DICOM is a specialized format used exclusively for medical imaging. It allows the storage and transmission of medical images along with metadata such as patient information, imaging modality, and image acquisition parameters. DICOM ensures that medical images are standardized and interoperable across different systems and devices.

- **GIF (Graphics Interchange Format)**: Primarily used for simple images and animations, GIF uses indexed color and lossless compression. Due to its limited color palette, it is not suitable for high-fidelity applications like medical imaging.

In medical contexts, the DICOM format is the most prevalent, as it allows not only the image data to be stored but also patient and procedure information, ensuring proper medical record-keeping and interoperability between imaging systems[36]. High-quality, lossless formats such as TIFF and PNG may also be used in situations where image fidelity is paramount [6]. Table 1 summarizes the key characteristics of common image file extensions.

Table 1. Properties of some image extensions

Feature	JPEG	PNG	TIFF	BMP	DICOM	GIF
Lossy or Lossless	Lossy	Lossless	Lossless	Lossless	Lossless	Lossless
Compression	High	Moderate	Moderate	Low	Low	High
Color Depth	24-bit	24-bit, 32-bit	24-bit, 32-bit, 48-bit, etc.	24-bit	16-bit	8-bit
Transparency	Yes (limited)	Yes	Yes	No	No	Yes
Animation	No	No	No	No	No	Yes
Common Use Cases	General-purpose images, photography	High-quality images, logos, transparency	High-quality images, medical imaging, scientific data	Simple images, older applications	Medical imaging	Simple animations, graphics

II.2.5. Modalities of Medical Images

Medical images are specialized types of digital images used in healthcare to visualize the anatomy and function of the human body. These images are crucial for diagnosing diseases, planning treatments, and conducting surgeries. Various medical imaging modalities include:

- **X-rays**: are primarily used to visualize bones, joints, and other dense tissues. They work by passing X-rays through the body, with denser tissues absorbing more radiation than softer tissues. This creates a black and white image where bones appear white and softer tissues appear darker. X-rays are commonly used to diagnose fractures, dislocations, arthritis, and certain types of cancer[30].

- **Magnetic Resonance Imaging (MRI)** is a powerful imaging technique that produces detailed images of soft tissues, organs, and blood vessels. It uses a strong magnetic field and radio waves to align the body's hydrogen atoms. The signals emitted by these

atoms are used to create high-resolution images. MRI is commonly used to diagnose conditions affecting the brain, spine, muscles, joints, and blood vessels[38] .

- **Computed Tomography (CT)** is a non-invasive imaging technique that generates cross-sectional images of the body. It uses X-rays and computer algorithms to create 3D images from multiple angles. CT scans are used to diagnose injuries, tumors, infections, and other conditions affecting various body parts[38].

- **Ultrasound** is an imaging technique that uses sound waves to create images of soft tissues. High-frequency sound waves are sent into the body, and the echoes are used to create images. Ultrasound is commonly used in prenatal care to monitor the development of the fetus, as well as to examine the heart, liver, kidneys, and other organs[38].

- **Positron Emission Tomography (PET)** is a nuclear medicine imaging technique that generates 3D images of metabolic activity in the body. A radioactive tracer is injected into the body, and the PET scanner detects the gamma rays emitted by the tracer. PET scans are used to diagnose cancer, brain disorders, and heart conditions[38].

Medical images are typically in grayscale (Fig. 3), where the pixel intensities represent varying levels of density or material composition in the human body [14]. These images require high precision and fidelity because even minute details can be critical for accurate diagnosis [15].

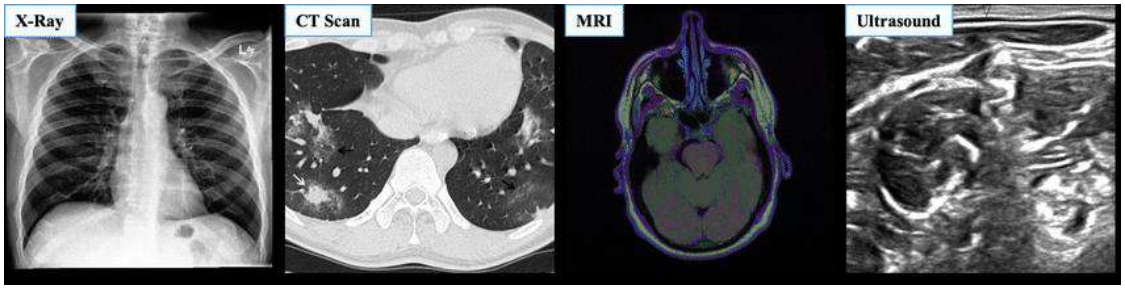


Fig 3.Sample images from various medical imaging modalities[39].

II.3. **Introduction to Image Watermarking**

II.3.1. What is Image watermarking ?

Image watermarking is a technique used to embed additional information into an image in a way that does not significantly affect its visual quality [40]. This embedded data, called the "watermark," can be used for various purposes, such as copyright protection, verification of ownership, and maintaining the integrity of digital media [41]. Watermarking has become an essential tool in safeguarding intellectual property in an era where digital content is widely shared and easily distributed across different platforms [39].

The demand for robust and secure watermarking techniques has increased with the proliferation of digital content and the need to protect sensitive information, especially in fields like medical imaging [42]. By embedding a watermark, the origin and authenticity of the image can be verified, even after undergoing various transformations, such as compression or noise addition [35].

II.3.2. Applications of Image Watermarking

Image watermarking is a versatile technique with numerous applications across various fields. Here, we'll explore several key areas where image watermarking plays a crucial role:

a) Copyright Protection

Watermarking is commonly used to protect digital content from unauthorized use or reproduction. By embedding a watermark that identifies the copyright holder or the origin of the image, creators and owners can assert their rights and prevent misuse of their intellectual property.

Example: Photographers and artists embed watermarks in their digital photos to prevent unauthorized sharing and to ensure that their work is credited to them.

b) Digital Rights Management (DRM)

In the context of DRM, watermarking helps in managing and enforcing usage rights. It ensures that digital media, such as movies, music, and e-books, is distributed and accessed according to licensing agreements.

Example: Streaming platforms use watermarking to track and manage the distribution of their media content, ensuring that it is not illegally copied or shared.

c) Medical Imaging

Watermarking in medical imaging ensures the authenticity and integrity of medical images, such as X-rays, MRIs, and CT scans. This is crucial for maintaining accurate diagnoses and preventing tampering with patient data.

Example: Hospitals and diagnostic centers use watermarking to secure medical images transmitted over networks, ensuring that they remain unaltered and can be traced back to the source.

d) Forensic Analysis

Watermarking is used in forensic analysis to track the provenance of images and verify their authenticity. It helps in identifying and mitigating instances of tampering or forgery.

Example: Law enforcement agencies use watermarking to ensure the integrity of evidence images, such as crime scene photographs, to support their validity in court.

e) Branding and Marketing

Companies use watermarking to protect their brand assets and to promote their identity. By embedding logos or brand marks into promotional images, they can prevent unauthorized use and enhance brand recognition.

Example: Businesses add watermarks to their advertising images to ensure that their content is not misused by competitors or third parties.

f) Document Authentication

Watermarking is used to authenticate and verify the legitimacy of important documents, including certificates, diplomas, and contracts. It helps prevent counterfeiting and ensures that the document's source is credible.

Example: Educational institutions embed watermarks in diplomas to prevent forgery and to verify the authenticity of the issued certificates.

g) Secure Communication

In secure communication systems, watermarking can be used to embed information or tags within images to facilitate secure transmission and prevent unauthorized access or modification.

Example: Secure communication platforms use watermarking to embed cryptographic keys or authentication tokens within transmitted images, ensuring the security and integrity of the communication.

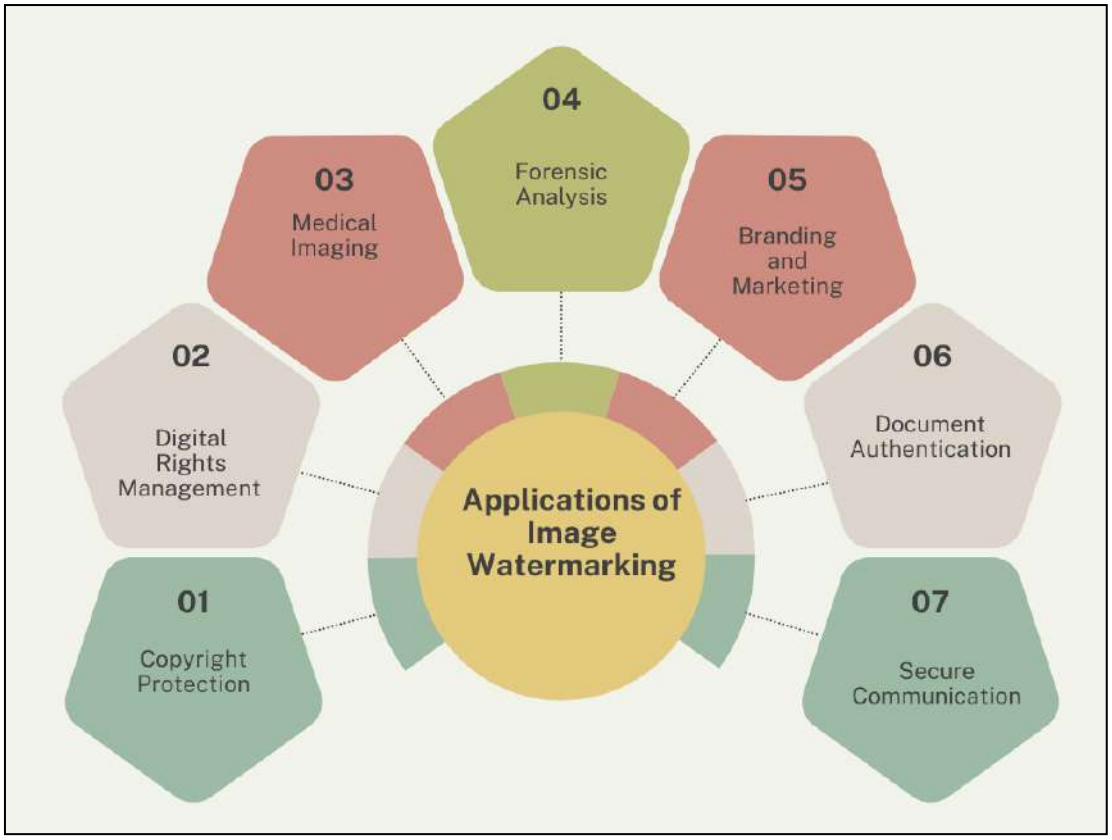


Fig 4. Application of watermarking

II.3.3. Categorization of watermarking schemes

Watermarking can be classified in multiple ways based on different factors: the visibility of the watermark, the robustness against attacks, blindness and reversibility. Each type has specific use cases and challenges.

a) Perceptibility:

- *Visible*: A watermark that is deliberately visible on the image, such as a logo or text overlay, and does not require extraction techniques to detect [6]. It is mainly used for branding and copyright purposes.
- *Invisible*: The watermark is embedded in such a way that it is imperceptible to the human eye, ensuring the original image's visual quality is preserved[43]. Invisible watermarks are used for security and authentication purposes, and extraction methods are required to retrieve them.

b) Robustness:

- *Fragile*: Fragile watermarks are sensitive to even slight modifications of the image. They break or degrade if any tampering occurs, making them useful for integrity verification [44].
- *Robust*: Designed to withstand various image processing attacks (e.g., compression, noise, resizing) [45] Robust watermarks are useful for copyright protection, ensuring the watermark remains intact under common manipulations [46].

c) Blindness :

- *Blind*: The extraction of the watermark can be performed without the need for the original image [47]. This method is highly desirable for real-world applications, such as medical image watermarking, where the original image might not be accessible [48].
- *Non-Blind*: Requires the original image to extract the watermark. This method provides more robust watermarking but at the cost of needing both the original and the watermarked images for verification [47].

d) Reversibility:

- *Reversible*: Allows the original image to be completely restored after the watermark has been extracted [49]. This technique is crucial in applications where preserving the exact original content is essential, such as medical or satellite imagery [50].

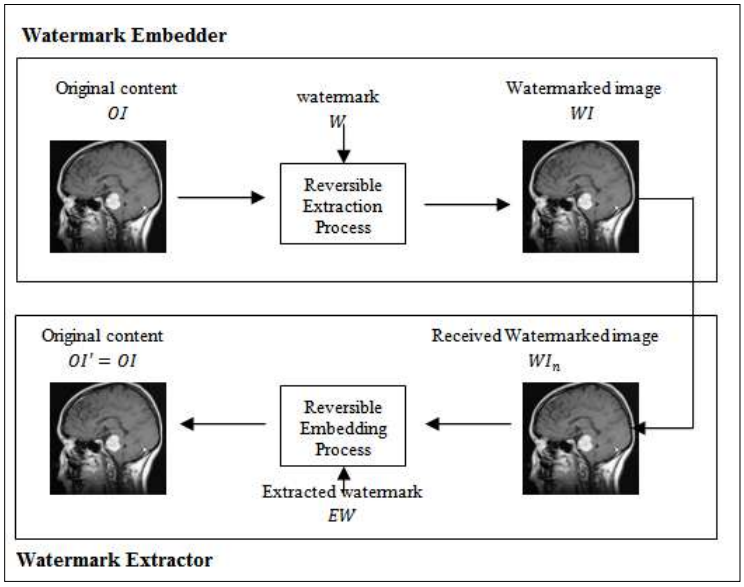


Fig 5. Embedding and Extration Process of Reversibe Watermarking

- *Irreversible:* The watermark is embedded in such a way that the original image cannot be perfectly restored once the watermark is removed [51]. This is acceptable in contexts where minor distortions in the image are not critical, such as entertainment media [49].

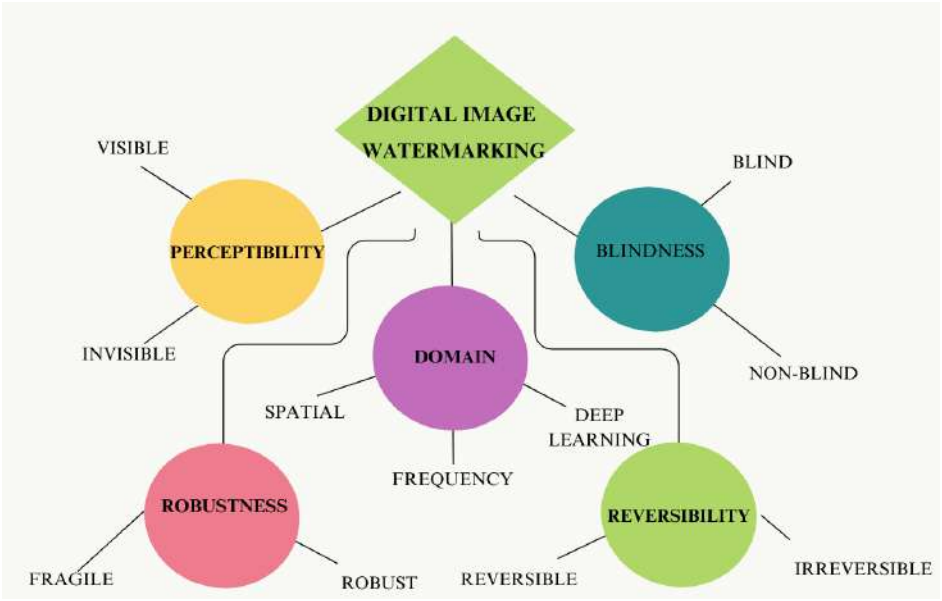


Fig 6. Classification of digital image watermarking

II.3.4. Watermark Embedding and Extraction Process

Watermarking typically involves two critical stages: embedding the watermark into the host image and extracting the watermark from the watermarked image. These processes are fundamental in understanding how watermarking works in practice.

a) Watermark Embedding Process

The watermark embedding process refers to the integration of additional data (the watermark) into the host image. The challenge is to embed the watermark in a way that it is imperceptible to human eyes, while still maintaining a high level of robustness against various image manipulations like compression, resizing, and attacks.

- **Step 1:** *Preprocessing the Watermark*

  Before embedding, the watermark (which could be a binary image, a logo, or text) is often processed to optimize it for the embedding process. Depending on the technique used, this may involve transforming the watermark into a different domain, such as the frequency domain using a Discrete Wavelet Transform (DWT) [40].

- **Step 2:** *Selecting Embedding Location*

The watermark is embedded in either the spatial domain (modifying pixel values directly) or the transform domain (modifying the transformed coefficients of the image). Transform-domain methods like DWT or Discrete Cosine Transform (DCT) provide greater robustness against attacks [33].

- **Step 3:** *Embedding the Watermark*

The actual embedding process involves modifying specific pixels or coefficients in the host image based on the watermark's data. The strength of these modifications needs to balance between imperceptibility (to avoid visual degradation) and robustness (to withstand various attacks) [27]. The formula for watermark embedding can be generalized as:

$$WI = HI + \alpha \cdot W \qquad (1)$$

Where *HI* is the host image, *W* is the watermark, α is a scaling factor, and *WI* is the watermarked image [28].

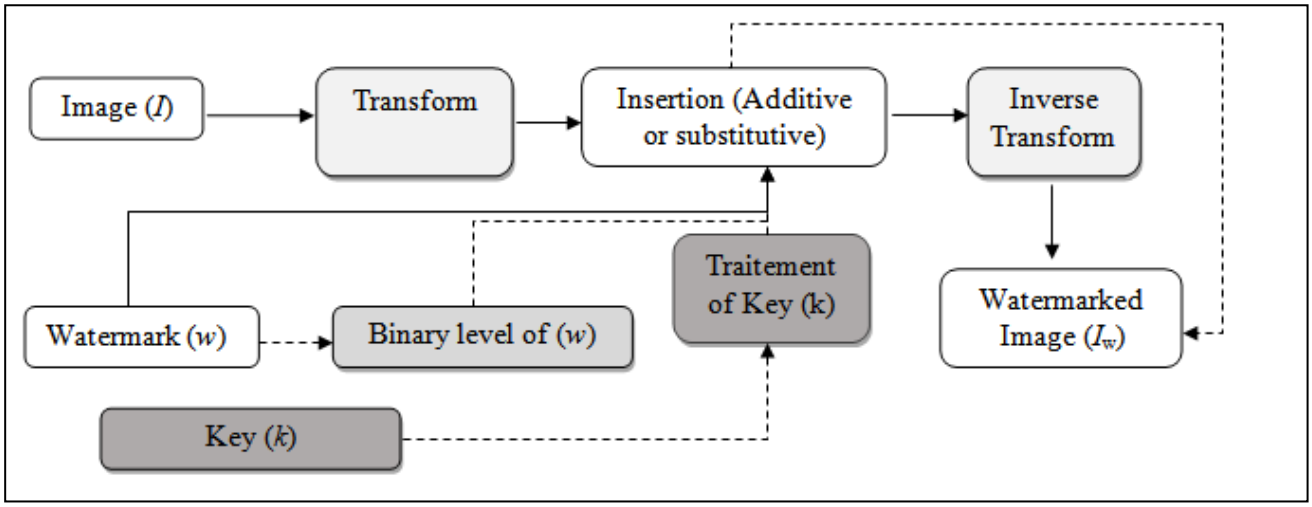


Fig 7.schema of watermark embedding process

b) Watermark Extraction Process

The watermark extraction process retrieves the watermark from a potentially tampered watermarked image. This process should be robust enough to detect the embedded watermark, even if the watermarked image has undergone attacks or modifications.

Step 1: *Preprocessing the Watermarked Image*

Before extraction, the watermarked image is preprocessed in the same domain where the watermark was embedded. If the watermark was embedded in the frequency domain, the same transformation (e.g., DWT or DCT) is applied to the watermarked image [17].

- **Step 2***: Extraction Algorithm*

The extraction algorithm retrieves the embedded watermark from the processed image. The method of extraction varies based on the embedding technique, but it often involves comparing the modified image data to the original image data (if available) or using a correlation-based approach to detect the presence of the watermark [52].

- **Step 3:** *Post-processing*

After extraction, the watermark may undergo post-processing to reconstruct it in a form close to the original. This is especially important if the watermark has been distorted by attacks. Metrics like Normalized Correlation (NC) or Bit Error Ratio (BER) are used to evaluate the fidelity of the extracted watermark[53] .

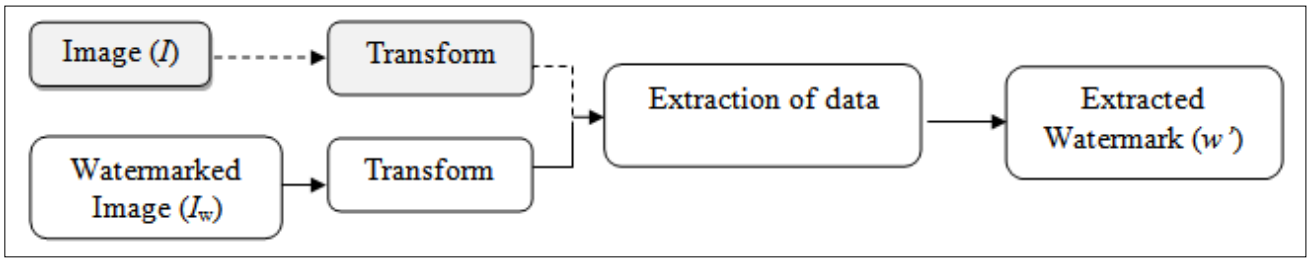


Fig 8. Schema of watermark extraction process

II.3.5. Watermarking techniques

When discussing watermarking techniques, it's essential to categorize the methods into different domains and approaches. Here's an organized breakdown of watermarking techniques:

a) Handcrafted techniques

➢ *Spatial domain watermarking techniques*

- Overview**:** In this technique, the watermark is directly embedded into the pixel values of the image. Spatial domain methods are generally simple and fast but tend to be less robust against attacks.
- Example Techniques:
  - Least Significant Bit (LSB): The watermark is embedded into the least significant bits of the pixel values. This method is easy to implement and computationally efficient but is vulnerable to noise, compression, and other image manipulations [54].
  - Additive Watermarking: The watermark is added to the pixel values, usually using an alpha blending factor. This method offers slightly better robustness than LSB [47].
- Pros**:** Fast and easy to implement.
- Cons**:** Low robustness, especially against compression and noise.

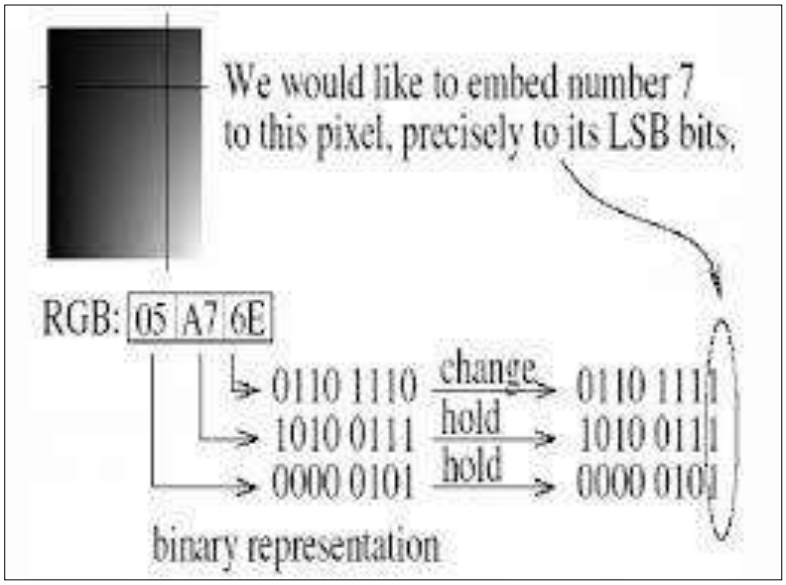


Fig 9.. Least Significant Bit (L.S.B) Techniques of Color Image Watermarking.

➢ *Transform domain watermarking techniques*

- Overview**:** In transform domain methods, the image is transformed into a different domain (e.g., frequency) before the watermark is embedded. These techniques are more robust to attacks such as compression and noise than spatial domain techniques.

- Example Techniques:

A. *Discrete Cosine Transform (DCT)*: The image is converted to the frequency domain using DCT, and the watermark is embedded in the transformed coefficients. DCT-based methods are widely used in JPEG compression and offer good robustness [55].

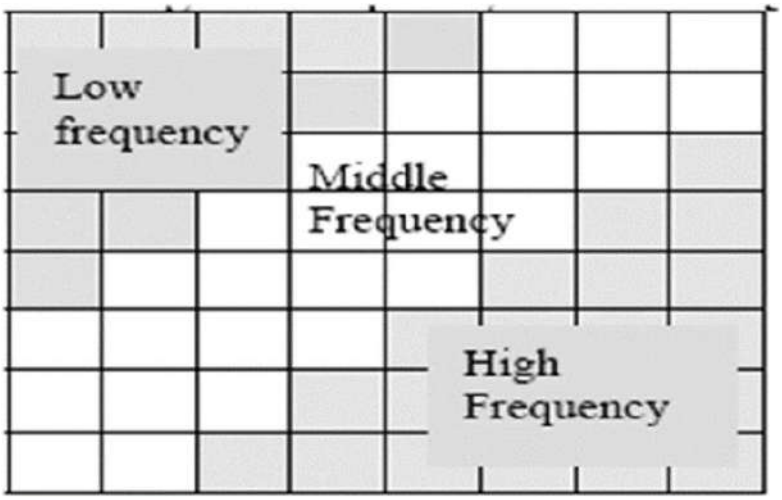


Fig 10.Coefficient matrix of DCT

B. *Discrete Wavelet Transform (DWT)*: DWT decomposes the image into multi-resolution frequency bands. The watermark can be embedded in different sub-bands, providing a good balance between imperceptibility and robustness. DWT is effective against attacks like compression, resizing, and noise [56].

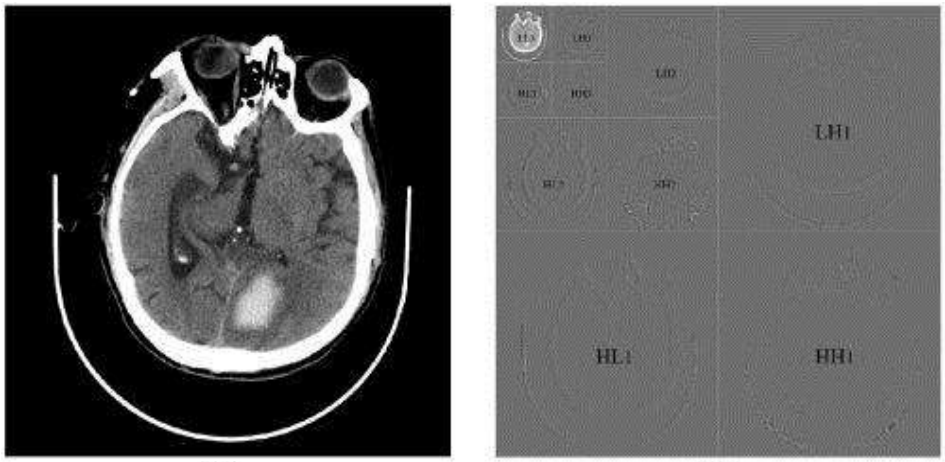


Fig 11. Three-level-discrete-wavelet-decomposition-of-the-host-medical-image-ahost-CT-image

C. *Singular Value Decomposition (SVD)*: SVD transforms the image into singular values, and the watermark is embedded in the singular matrix. SVD offers high robustness, especially against geometric transformations [43].

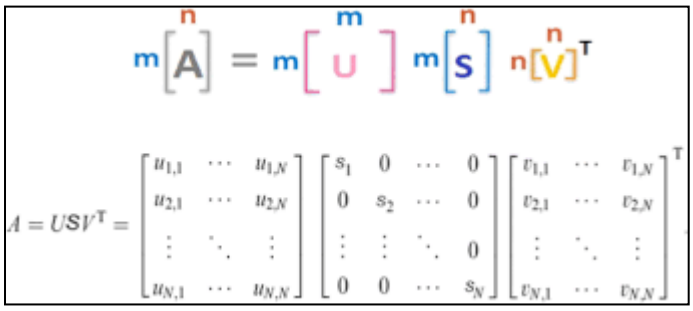


Fig 12. SVD decomposition[57]

- Pros**:** Better robustness, especially against common image processing attacks e.g., JPEG compression).

- Cons**:** More computationally complex than spatial domain methods.

➢ *Hybrid Handcrafted Watermarking Techniques*

- Overview**:** Hybrid techniques combine the advantages of multiple methods to improve both imperceptibility and robustness. For instance, they may embed a watermark using both spatial and transform domain techniques.

- Example Techniques:
    - LBP-DWT (Local Binary Patterns - Discrete Wavelet Transform): This method combines the local texture feature extraction of LBP with the frequency decomposition of DWT. This combination enhances robustness and imperceptibility, making it particularly suitable for medical images where preserving quality is critical[11].
    - DWT-SVD**:** A common hybrid method where DWT is used for multi-resolution analysis, and SVD provides robustness to transformations. This method is highly effective against noise, cropping, and compression attacks [31].
- Pros**:** Offers high robustness and imperceptibility.
- Cons**:** Complex to implement and may require more computational power.

b) Deep Learning-Based Watermarking Techniques

Deep learning is a subset of machine learning that uses artificial neural networks with multiple layers to learn complex patterns in data. It's inspired by the structure and function of the human brain, where information is processed through interconnected layers of neurons [39].Fig. 13 shows the basics of DNN.

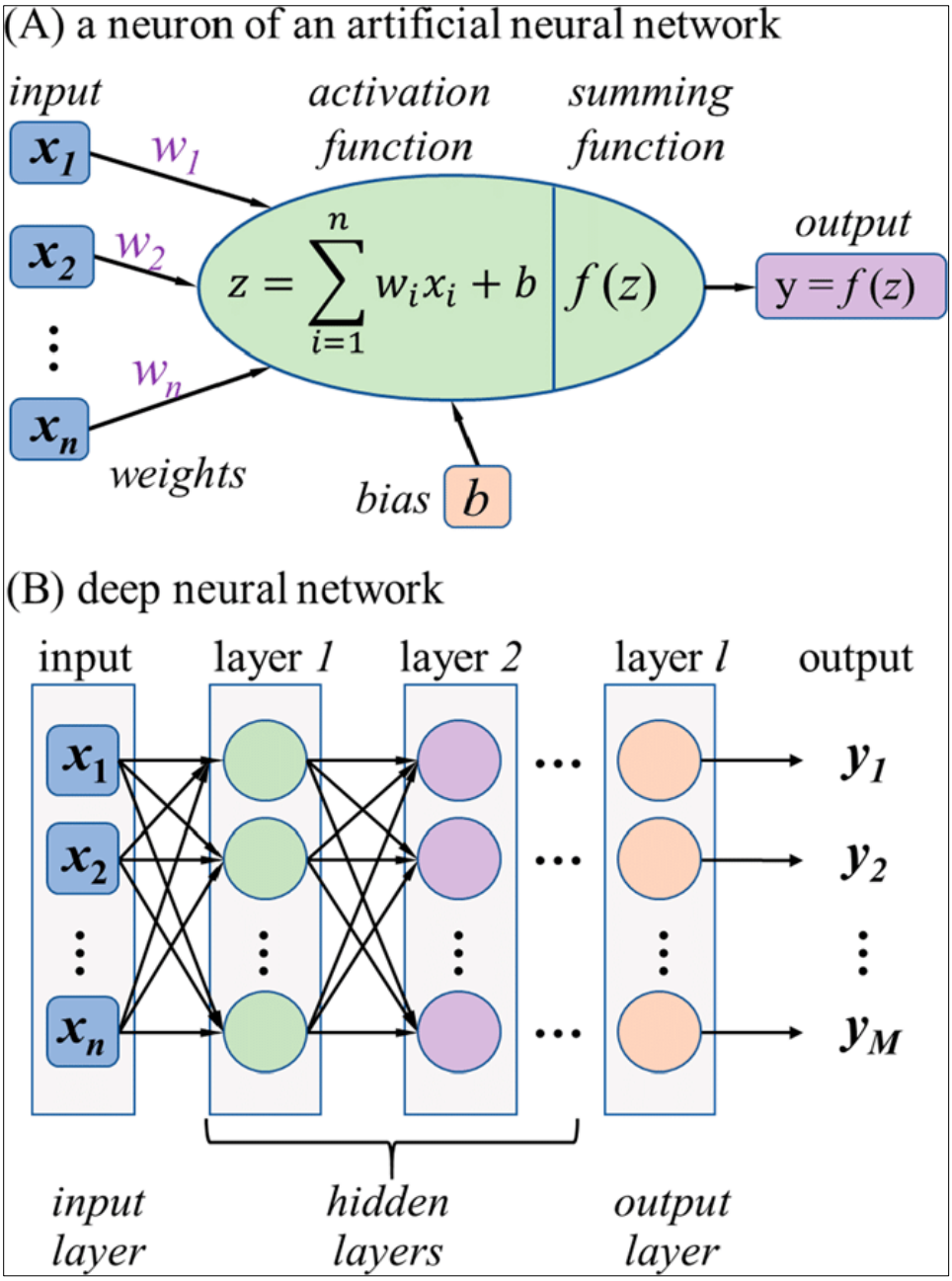


Fig 13. Basics of Deep Neural Network[58]

Deep learning models can automatically learn features from raw data, making them highly effective for tasks like image recognition, natural language processing, and speech recognition[59]. The depth of the neural network, combined with powerful algorithms and large datasets, enables deep learning models to achieve state-of-the-art performance in various domains.

With the rise of deep learning, new techniques have emerged that use neural networks:

- Deep Neural Networks (DNNs): A type of artificial neural network with multiple hidden layers, allowing it to learn complex patterns in data. DNNs are highly effective for tasks like image recognition and natural language processing. Each neuron in a DNN processes inputs from the previous layer, passes the result to the next layer, and learns to represent increasingly abstract patterns. To effectively train a DNN for watermarking tasks, consider designing a suitable architecture, training the DNN on the data, and evaluating its performance. Key considerations include data quality, hyper parameter tuning, regularization, and potential use of transfer learning[58] .

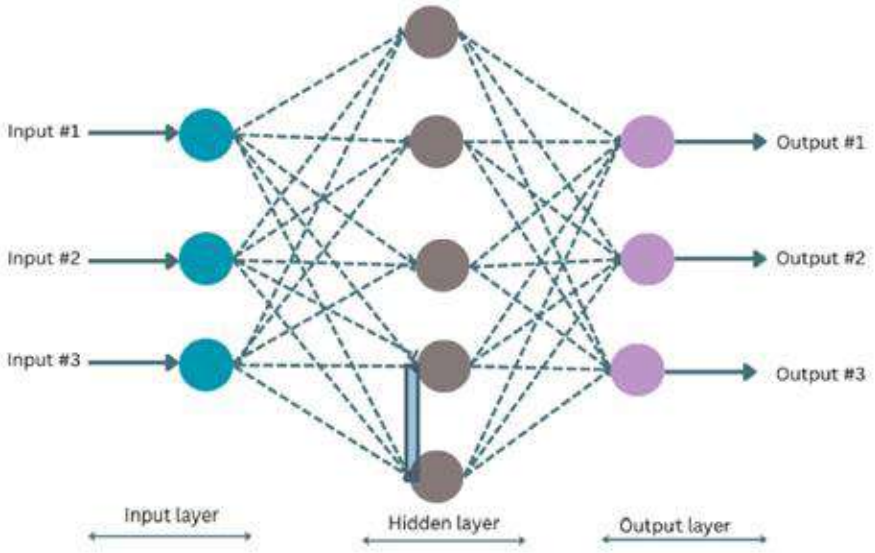


Fig 14. DNN architecture[60]

- Convolutional Neural Networks (CNNs): A type of deep neural network specifically designed for processing grid-like data, such as images. CNNs consist of convolutional layers that extract features, pooling layers that downsample feature maps, and fully connected layers for classification. They have been successfully applied to tasks like image classification, object detection, and natural language processing[61].

Example Techniques:

- CNN-based Watermarking: CNNs can automatically learn features from the image and embed the watermark in a way that maximizes robustness and imperceptibility. The network can be trained to adapt to various attack scenarios, making it highly flexible[8].

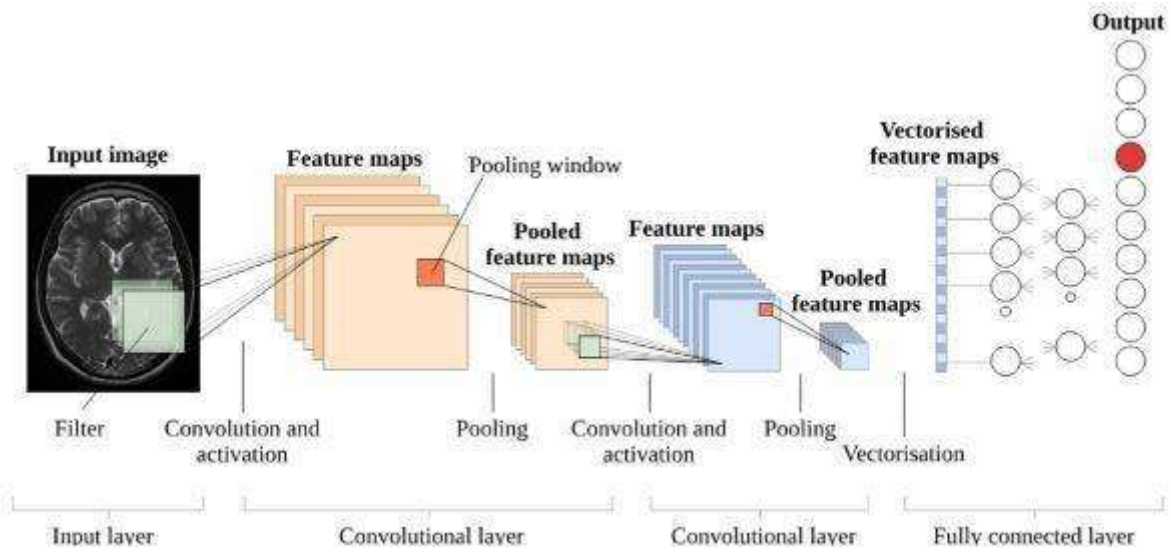


Fig 15.Architecture of convolutional neural network [44]

- Pre-trained Models**:** These are deep learning models that have been trained on large datasets to perform specific tasks. They can save time and resources compared to training models from scratch and often achieve better performance. Examples include image recognition, natural language processing, and speech recognition models. Pre-trained CNNs, such as VGG16 [62]or ResNet[24], can be fine-tuned for watermark embedding and extraction tasks. These models leverage transfer learning to achieve good performance even with limited training data[9], [63].

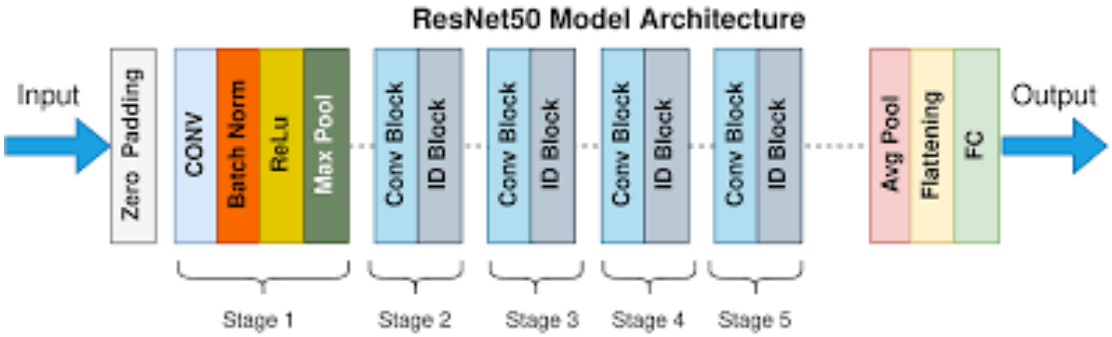


Fig 16.ResNet50 Architecture[64]

- Autoencoders: A type of artificial neural network that learns to compress and reconstruct data. It consists of two main components: an encoder that maps the input

data to a lower-dimensional representation (latent space), and a decoder that reconstructs the original data from the latent representation. The goal of an autoencoder is to learn a compressed representation of the data that captures the most important features while discarding irrelevant information. This can be useful for tasks such as dimensionality reduction, denoising, and anomaly detection. The watermark can be embedded by modifying features in the latent space, and it can be extracted by extracting features from the watermarked content [65].

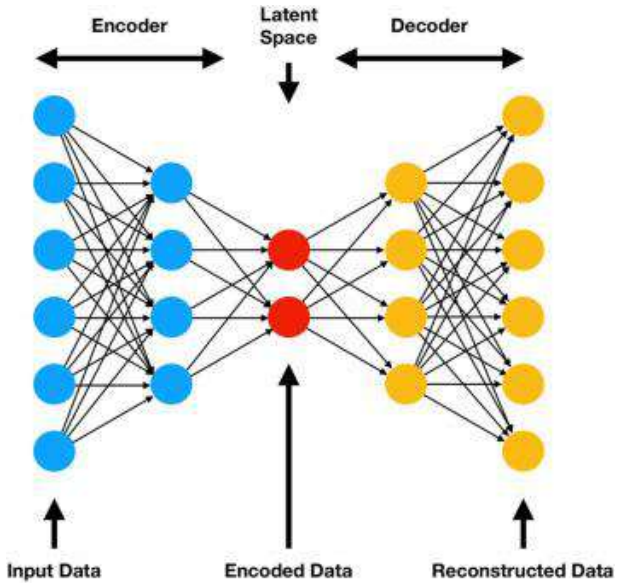


Fig 17. Auto encoder architecture[66]

II.3.6. Watermarking's performance metrics

When discussing watermarking, it is essential to highlight the key properties that determine the effectiveness and utility of a watermarking system. These properties help in evaluating the performance of different watermarking techniques and guide the selection of appropriate methods based on the specific use case, such as protecting medical images or multimedia content. The main properties include perceptibility, robustness, capacity, and blindness.

a) Perceptibility (Imperceptibility):

Perceptibility refers to the degree to which the embedded watermark affects the visual quality of the watermarked image. A watermark should be imperceptible, meaning the presence of the watermark should not degrade the quality of the image or be easily noticed by the human eye.

- Importance**:** In applications like medical imaging, maintaining the visual fidelity of the image is crucial since doctors rely on these images for diagnosis. Any visible degradation can lead to misinterpretation [46].
- Measurement**:** Common metrics for evaluating perceptibility include:

➢ Peak Signal-to-Noise Ratio (PSNR): Higher PSNR indicates better imperceptibility.

$$\mathrm{PSNR_{dB}} = 10\log_{10}[\,\mathrm{N} * \mathrm{M}\,\frac{\max \mathrm{I}^2(\mathrm{i,j})}{\sum \mathrm{i,j}\,[\,\mathrm{I(i,j)} - \mathrm{J(i,j)}\,]^2}\,] \qquad (2)$$

Here, I and J represent the original image and the watermarked image, respectively, both having sizes N × M.

➢ Structural Similarity Index (SSIM): SSIM measures the structural similarity between the original and the watermarked image [67].

$$\mathrm{SSIM(x,y)} = \frac{(2\mu_x\,\mu_y + C_1)(2\delta_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\delta_x^2 + \delta_y^2 + C_2)} \qquad (3)$$

The symbols represent the following: μ_x : The average of x. μ_y: The average of y. δ_x^2: The variance of x. δ_y^2: Variance of y. δxy: Covariance between x and y.

- Challenges**:** Balancing imperceptibility with robustness is often difficult, as stronger watermarks are more visible but more resistant to attacks, while less noticeable watermarks are easier to remove [67].

b) Robustness:

Robustness refers to the watermark's ability to resist various attacks or modifications applied to the watermarked image. The watermark should remain detectable even after the image has undergone common manipulations such as compression, resizing, noise addition, cropping, or geometric transformations.

• Importance**:** For applications like content authentication, copyright protection, or medical data integrity, robustness is critical. The watermark needs to survive even after multiple image processing operations or attacks [5].

• Measurement**:** Common metrics for evaluating robustness include:

- A Normalized Correlation denoted as NC , indicates how similar the original image and the one under attack are. Two images are highly identical if the coefficient (NC) is 1. if the coefficient (NC) is 0, two images are completely different [62]. This coefficient ranges from 0 to 1. Eq. 4 is utilized to calculate this metric:

$$NC(w, w_e) = \frac{\sum_{p=1}^{M} \sum_{q=15}^{N} (w(i,j) - \overline{w}\,(i,q)(w(i,j) - \overline{w_e}\,).}{\sqrt{(\sum_{p=1}^{M} \sum_{q=10}^{N} w(i,j) - \overline{w}^2)\ (\sum_{p=1}^{M} \sum_{q=1}^{N} w_e(i,j) - \overline{w_e}^2)}} \quad (4)$$

Where $w(p, q)$ and $w_e(p, q)$ denote the intensities of the pixel at coordinates (p, q) in the original w and the extracted watermark w_e, respectively. The averages of intensities for the watermark image *w* and the extracted watermark w_e are denoted by $\overline{w}$ and $\overline{w_e}$, respectively.

- The Bit Error Rate (BER) assesses the number of errors between the extracted watermark and its original, while the Bit Correct Rate (BCR) assesses the number of corrects between the extracted watermark and its original [31], [67]. The following expression for its estimation is given by Eq. 5 :

$$BCR = 1 - BER = \frac{\text{Number of correctly decoded bits}}{\text{Total number of embedded bits}} \quad (5)$$

c) Capacity:

Capacity refers to the amount of data that can be embedded into the image without affecting its quality or robustness. This property is measured by how many bits of the watermark can be stored within the image.

• Importance**:** Higher capacity allows embedding more detailed information, such as multiple watermarks or larger watermarks, into the image. In some cases, capacity is crucial when dealing with complex content authentication or copyright information [68].

- Challenges**:** Increasing capacity can often reduce imperceptibility or robustness. The balance between capacity, imperceptibility, and robustness must be carefully managed based on the application's requirements [69].

Each watermarking system should be designed with the intended application in mind, considering the necessary balance between these properties. For example, in medical imaging, perceptibility is prioritized to ensure diagnostic accuracy, while in copyright protection, robustness is more critical.

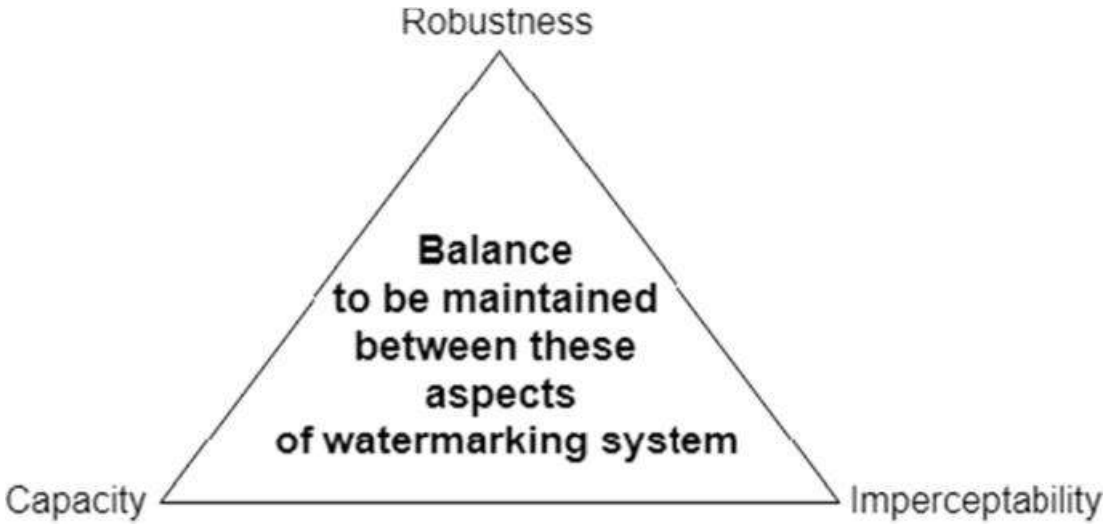


Fig 18.Relationship between major parameters of watermarking[70]

II.3.7. Watermarking attacks

Watermarking attacks are strategies used to degrade or remove the embedded watermark from an image. These attacks can be broadly categorized into common image processing attacks and geometric attacks. Understanding these attacks is essential for developing robust watermarking methods that can withstand various forms of manipulation.

a) Common Image Processing Attacks

- **Noise Addition**
  - Description: Noise addition involves introducing random variations in pixel values, which can obscure or distort the watermark. Common types of noise include salt-and-pepper noise and Gaussian noise.

- Salt-and-Pepper Noise: Randomly adds white and black pixels to the image, simulating salt-and-pepper-like speckles.
- Gaussian Noise: Adds variations according to a Gaussian distribution, affecting the overall image quality. The Gaussian function for noise addition is typically defined as:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2-y^2}{\sigma^2}} \quad (6)$$

- Impact on Watermarking: Noise can interfere with the watermark's visibility and make it harder to detect or extract, especially if the watermarking method is not designed to be noise-resilient.
- Example: Adding Gaussian noise to an image can cause subtle distortions that might render a watermark ineffective or less discernible.

➢ **Filtering**

- Description: Filtering involves applying various types of image filters that alter the image's appearance and can affect the watermark. Common filters include:
  - Median Filter: Removes noise by replacing each pixel's value with the median value of its neighbors.
  - Gaussian Filter: Blurs the image by averaging pixel values in a Gaussian distribution.
  - Average Filter: Averages the pixel values in a neighborhood to smooth the image.
  - Sharpening Filter: Enhances edges and details by emphasizing differences between neighboring pixels.
- Impact on Watermarking: Filtering can distort or blur the watermark, potentially reducing its visibility or integrity.
- Example: Applying a Gaussian filter can blur both the image and the watermark, making the watermark harder to distinguish.

- **Compression**

  - Description: Compression techniques reduce the image file size but can alter pixel values and degrade watermark quality.
    - JPEG Compression: A lossy compression method that can introduce artifacts and reduce image quality.
    - PNG Compression: A lossless compression method that maintains image quality but may still affect watermark visibility.
  - Impact on Watermarking: Lossy compression, like JPEG, can significantly degrade the watermark, while lossless compression, like PNG, may have a lesser effect.
  - Example: Saving an image with JPEG compression can introduce blocky artifacts that distort the watermark.

- **Histogram Equalization**

  - Description: Histogram equalization adjusts the contrast of the image by spreading out the pixel intensity values. This can enhance or suppress certain features in the image.
  - Impact on Watermarking: This process can change the watermark's visibility by altering the image's pixel distribution.
  - Example: Applying histogram equalization to a watermarked image might make the watermark more or less visible, depending on how it affects the pixel intensity distribution.

- **Bit-Plane Removal**

  - Description: Bit-plane removal involves deleting or modifying specific bit planes of the image, affecting the precision of pixel values.
  - Impact on Watermarking: This can result in the loss of watermark information, as important bits containing the watermark data may be removed.
  - Example: Removing the least significant bit plane of a watermarked image can reduce watermark fidelity.

- **Gamma Correction**

  - Description: Gamma correction adjusts the brightness and contrast of the image based on a nonlinear transformation. expressed as:

$$I_{out} = I_{in}^{\gamma} \quad (7)$$

Where γ controls the brightness.

- Impact on Watermarking: Modifies the watermark's appearance, potentially making it less detectable.
- Example: Applying gamma correction to an image can enhance or diminish the visibility of the watermark depending on the gamma value used.

b) Geometric Attacks

➢ **Rotation**

- Description: Rotation involves turning the image by a certain angle. This can disrupt the alignment of the watermark with the image.
- Impact on Watermarking: Rotation can misalign the watermark, making it difficult to detect or extract accurately if the watermarking method is not rotation-invariant.
- Example: Rotating an image with a watermark by 45 degrees can result in a misaligned watermark that is challenging to recover.

➢ **Scaling**

- Description: Scaling changes the size of the image, either enlarging or reducing it.
- Impact on Watermarking: Scaling can stretch or compress the watermark, potentially making it less readable or visible. The watermark's effectiveness may depend on the scaling factor and the watermarking method's ability to handle such transformations.
- Example: Upscaling an image can spread out the watermark, reducing its contrast and making it harder to discern.

➢ **Cropping**

- Description: Cropping involves removing parts of the image, which can affect the areas where the watermark is embedded.
- Impact on Watermarking: Cropping can eliminate the watermark entirely if it is located in the cropped section. It can also reduce the watermark's effectiveness if only a portion of it is retained.
- Example: Cropping an image to remove a watermark embedded in the corner can completely erase the watermark if it was located in the cropped area.

Understanding these attacks helps in designing watermarking techniques that can better withstand various forms of manipulation. Effective watermarking systems should

incorporate mechanisms to resist common image processing and geometric attacks, ensuring that the watermark remains detectable and intact under different conditions.

Fig. 20 shows various attacks on MRI image.

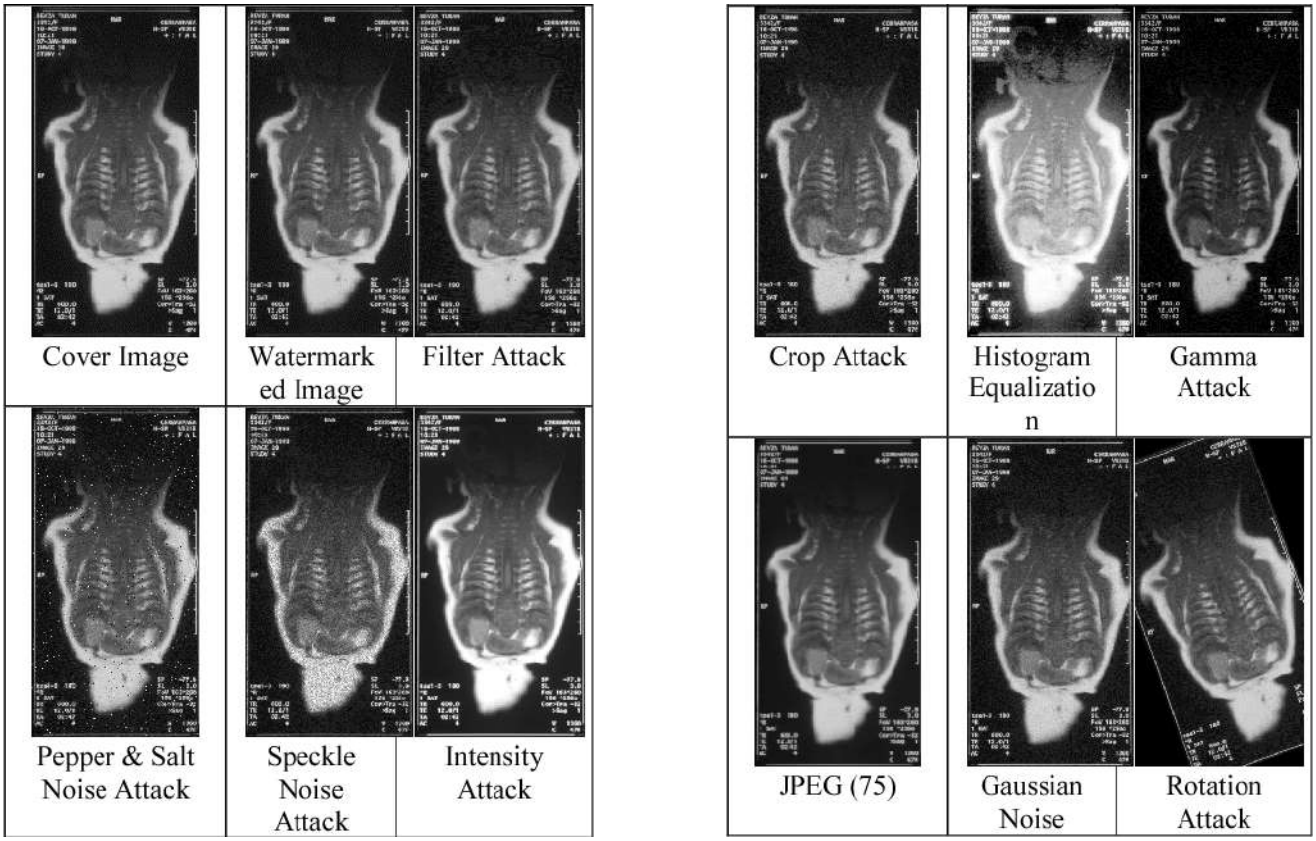


Fig 19.Watermarked MRI images after attacks[71]

III. Chapter 2: Image Watermarking: A State of the Art Review

III.1. Introduction

Image watermarking is a vital technique for protecting the intellectual property and integrity of medical images. It involves embedding a hidden watermark signal into the image, allowing for subsequent identification, authentication, and copyright protection. Traditional handcrafted watermarking techniques have been widely used, but recent advancements in deep learning have opened up new possibilities for improving the performance and robustness of watermarking systems.

One of the critical elements in medical image watermarking is feature extraction, a process that plays a pivotal role in embedding and extracting watermarks without compromising the quality of the images. This chapter explores the state-of-the-art methods in medical image watermarking, with a particular focus on comparing handcrafted feature extraction techniques to those derived from deep learning models. We will explore the principles and techniques underlying both approaches, analyze their advantages and limitations, and discuss their applications in the context of medical image security.

By understanding the strengths and weaknesses of each method, researchers and practitioners can make informed decisions about the most suitable approach for their specific needs. This chapter aims to provide a valuable resource for those working in the field of medical image security and watermarking.

III.2. Handcrafted Watermarking Techniques:

Handcrafted watermarking techniques play a crucial role in the realm of medical image protection. These methods often leverage various algorithms for embedding watermarks into images, relying on domain-specific features that enhance robustness against various attacks.

Some researchers disagree on classifying frequency transformations as methods capable of extracting features. However, we argue that they do extract valuable information, such as matrices, vectors, or properties that other techniques cannot. For instance, the Discrete Wavelet Transform (DWT) effectively captures vertical, horizontal, and diagonal edges, critical features that contribute to the image's overall structure and texture. This ability to

derive specific image characteristics underlines the significance of these handcrafted methods in watermarking applications.

In recent advancements, researchers have developed numerous hybrid and advanced watermarking techniques to improve both the security and robustness of medical image watermarking systems. These methods often involve a combination of frequency domain transformations, along with encryption strategies to enhance resistance against attacks. For instance, Alshanbari et al. [45] combined two different watermarking methods, integrating multiple watermarks into the medical image's region of non-interest (RONI). A SHA-256 secret key is first generated based on the image's region of interest (ROI), which is then compressed using the Lempel-Ziv-Welch (LZW) algorithm and combined with signature information to generate the final watermark. The watermark is embedded into the RONI, ensuring that the diagnostic content of the ROI remains unaltered, a crucial requirement for medical imaging applications **.**

Liu et al. [72]embeds a digital watermark through a two-step process: first, a private image is encoded using a fractal encoding method to generate specific encoding parameters that serve as the watermark. Next, these parameters are embedded into the discrete cosine transform (DCT) coefficients of the original image, allowing for reversible embedding. This dual encryption approach enhances robustness against attacks, enabling the extraction of the original private image from the watermarked carrier image without quality loss. Moreover,the Zear and Singh [73] method embeds a "Lump" image watermark and two text watermarks, "Signature" and "Symptoms," into a cover image. The "Lump" watermark is embedded by modifying singular values of a specific sub-band after applying DWT, DCT, and SVD. The text watermarks are embedded into different sub-bands using a simple formula. To extract the watermarks, the watermarked image undergoes similar steps, and the embedded watermarks are recovered by comparing singular values or modified coefficients. The "Lump" watermark is further processed using a BPNN to denoise it, while the text watermarks are decoded using Hamming decoding and arithmetic decoding.

The authors' method [56]involves creating a watermark composed of patient data and image acquisition parameters, which are converted into a QR code. The patient's photograph is compressed using the AMBTC algorithm, and the compressed image is encrypted using the RC4 algorithm with a key generated from the patient's information. To embed the watermark, the image is decomposed using the redundant discrete wavelet transform to extract the

frequency content of the image, and then the obtained LL subband coefficients are processed with Schur decomposition. The watermark bits are then integrated by modulating the successive eigenvalues obtained from the Schur decomposition. Finally, an inverse Schur decomposition is used to generate the updated coefficients, and an inverse redundant discrete wavelet transform is used to generate the watermarked image. Watermark extraction is a blind process that involves extracting the frequency content of the image using the discrete wavelet transform, then subjecting the resulting coefficients to Schur decomposition. Finally, the watermark bits are recovered from the eigenvalues. The recovered bits are used to create the QR code.

Additionally, this paper by Agrawal et al. [74] introduces a non-blind watermarking technique for medical images based on Discrete Cosine Transform (DCT). The method involves dividing the medical image into blocks and applying DCT to each block. The watermark, which is the Electronic Patient Record (EPR), is embedded into the middle frequency band of the DCT coefficients, known for its resistance to attacks. To improve the reliability of the embedded watermark, the EPR data is first encoded using Convolutional Error Correcting Code (ECC) and then decoded using Viterbi decoding. Additionally, M-ary modulation is applied to further enhance the watermark's robustness.

Singh et al. [75] propose a secure multiple watermarking method for medical images using DWT, DCT, and SVD. The method embeds both a medical image watermark (for identity authentication) and a text watermark (containing patient information) into a host medical image. In the embedding process, the host image is decomposed using DWT, and the LL subband is further processed with DCT and SVD. The watermark image is also transformed using DCT and SVD, and its singular values are embedded into the singular values of the host image. The text watermark, after being encrypted, is embedded into the HH subband of the host image. In this work, Sabbane et al. [28] present a novel region-based medical image watermarking technique. The method involves embedding numerical information (the watermark) into the original image. The key innovation is the use of polynomial transform to decompose the image into two parts: structure and texture. By focusing on the texture component, which contains less diagnostically relevant information, we can embed the watermark while maintaining high fidelity to the original medical image.

In another approach, Verma et al. [76] presented another hybrid watermarking technique combining DWT and SVD. The LL subband of the host image is chosen after DWT

decomposition, and SVD is applied to embed the watermark. This hybrid technique takes advantage of the low-frequency sub-band (LL) to ensure a high level of imperceptibility while maintaining robustness against various attacks. The combination of SVD and DWT provides a strong framework for embedding watermarks without significantly affecting the medical image's diagnostic quality .The proposed GWO-optimized hybrid DWT-DCT-SVD scheme in Hindang et al.'s [55] work involves embedding three binary watermark images one at a time into a cover image. GWO determines the optimal gain value for watermark embedding, while DWT and DCT select the appropriate frequency bands. SVD converts the DCT coefficients into a singular matrix for watermark insertion. This combination of transforms aims to achieve robustness, invisibility, and security in the watermarking process. Zairi et al.[77] present a fragile watermarking approach utilizing an optimal embedding technique combined with Huffman coding. The watermark image is encrypted using Huffman coding to create a scrambled version. The algorithm identifies the highest entropy value to select the most suitable non-overlapping blocks for embedding, utilizing the least significant bits (LSBs) of these blocks. This approach enhances safety and security through encrypted watermarking.

To address the crucial issue of authentication, Gull et al. [78] propose a fragile watermarking technique designed to detect and locate tampering in medical and general images. The cover image is divided into 4x4 non-overlapping pixel blocks, and each block is further divided into two 4x2 blocks: the Upper Half Block (UHB) and the Lower Half Block (LHB). Information embedded in the LHB enables tamper detection, while information embedded in the UHB facilitates tamper localization. Hernandez et al. [79] propose a robust and secure medical image watermarking method that combines invisible and zero watermarking techniques. The method embeds an encoded and encrypted watermark signal into the frequency domain of the medical image, establishing a link between clinical information, the medical image, and patient identity. The invisible watermarking component ensures that the medical image remains visually unchanged, while the zero watermarking stage authenticates the patient's identity based on the detector's response and the watermark information.

Jana et al. [80] propose a self-embedding fragile watermarking scheme for tamper detection and recovery in images. The method divides the image into 4x4 blocks and generates watermarks from the 6 most significant bits (MSBs) of each pixel. Blocks are

categorized as smooth or complex based on pixel similarity. For smooth blocks, the mean value is used as recovery data, while for complex blocks, AMBTC-compressed data is used. Recovery data is embedded into the 2 least significant bits (LSBs) of corresponding blocks. The scheme's efficiency is evaluated using Tamper Detection Rate (TDR) and Tamper Recovery Rate (TRR). In this paper, Sayeh et al. [81] propose a watermarking approach for patient identification and watermark integrity verification. The watermark consists of two parts: a patient information fingerprint and an encrypted patient photograph. The medical image is decomposed into four sub-bands using discrete wavelet transform, and the watermark bits are embedded by modulating the mid-frequency coefficients. This approach can accommodate both the patient's photography and fingerprint, and it can also be extended to include error-correcting codes. The coefficient modulation ensures that the watermark is embedded imperceptibly, resulting in a watermarked image that is visually similar to the original.

Expanding on these, Gangadhar et al. [34] use DWT and SVD with Particle Swarm Optimization, Similar to Verma et al., this method uses DWT and SVD on the LL sub-band, enhanced by a particle swarm optimization technique to determine optimal embedding locations .The Sing and Sing [82] algorithm uses a combination of NSCT, DCT, and SVD for watermarking, aiming for high capacity, robustness, and imperceptibility. It's a non-blind method, requiring the original cover image for watermark extraction. The algorithm pre-processes both cover and watermark images for accurate extraction. Medical images are used as covers, and EPR messages are embedded into selected sub-bands with chosen gain factors to balance imperceptibility, robustness, and capacity. Anand and Singh [31] propose a joint spatial-transform domain technique to embed a fused medical image watermark into the cover image, utilizing chaotic sequencing and SVD for encryption. This approach demonstrates greater robustness and excellent invisibility but requires verification of other security features like statistical analysis and key sensitivity.

Moreover, another technique of Wu et al. [83] introduce a reversible data hiding method for medical images that enhances contrast. They begin by segmenting the image background and identifying the primary grayscale values within the segmented region. By excluding these corresponding histogram bins from expansion during data hiding, they selectively enhance the contrast of the region of interest (ROI) in medical images. To minimize potential visual distortions caused by the pixel distribution, they propose a novel pre-processing strategy. Their method enables exact recovery of the original image from the enhanced image by embedding side information within it. In Swaraja's paper [84], a region-

based firefly algorithm combined with discrete wavelet transform (DWT) and Schur transforms is proposed for enhancing the security of medical images in telemedicine applications. The method focuses on ensuring authenticity of ownership and source origin during the exchange of medical images. By employing a blind watermarking technique, multiple robust watermarks are embedded in the region of non-interest (RONI) of the medical image using DWT-Schur. The effectiveness of the watermarking algorithm is evaluated based on its imperceptibility, robustness, and payload capacity, considering MRI, Ultrasound, and X-ray grayscale medical image modalities.

Here's a summary of comparative analysis of some handcrafted watermarking techniques for medical images in this table:

Table 2. Comparative Analysis of Some Classical Watermarking Techniques for Medical Images

Author	Watermark Type	Embedding Technique	Evaluation Metrics	Pros	Cons	Robustness
[56]	Fingerprint and patient photography	DWT, Shur Decomposition	PSNR(43.28) SSIM(0.9864) NC, BER	Employs AMBTC compression and RC4 encryption for watermark security.	Offers limited capacity and high computational complexity.	Robust
Singh et al. [75]	Logo	DCT, DWT, SVD	PSNR, BER, NC	Achieves enhanced robustness through combined transform schemes.	involves high computational complexity due to transform operations.	Robust
[73]	Lump image, Doctor sign, PI	DWT, DCT, SVD	PSNR(43.88), NC(0.9344), BER	Utilizes BPNN to reduce noise in the extracted watermark.	Requires significant computational resources due to transform-based operations.	Robust
Wu et al.[83]	PI	Background Segmentation, Contrast Enhancement	PSNR, SSIM, Capacity, Computational time	Enhances the contrast of the region of interest (ROI).	exhibits low embedding rates.	-
Gull et al.[78]	Logo	LSB Substitution	PSNR, SSIM, Capacity	Detects and localizes tampering at the 4x4 block level.	Does not explicitly consider watermark security measures.	Fragile
[84]	PI, Logo	Firefly Optimized Algorithm, DWT, Shur Transform	PSNR, Capacity, NC, BER	Incorporates multiple watermarks for enhanced security.	Has not been evaluated for error correction code compatibility.	Robust
[74]	EPR, Logo	DCT	PSNR, SSIM, NAE, BER, NCC	Addresses authentication, integrity, and copyright protection.	Lacks robustness against rotation attacks.	Robust

III.3. Handcrafted Techniques for Zero Watermarking:

Zero watermarking is a method where the watermark is not embedded directly into the image. Instead, a unique signature (watermark) is generated from the image's inherent features

and stored separately, allowing for watermark verification without modifying the host image. For a handcrafted zero watermarking method, the focus is on extracting distinctive features from the image using traditional techniques.

A variety of methods have been developed within this framework, such as Arevalo-Ancona and Cedillo-Hernández[85] propose a zero-watermarking scheme that utilizes K-means clustering for region of interest (ROI) detection and feature extraction. The K-means algorithm classifies data based on cluster proximity. By applying K-means clustering to the image, they identify ROIs containing important information. The discrete Fourier transform (DFT) is then applied to the ROI features, with high frequencies used to enhance robustness against geometric attacks. Additionally, Sobel edge detection is used to create a QR code watermark. This approach avoids watermark detection errors and improves system robustness. The master share is generated by performing an XOR operation on extracted ROI features and the watermark.

Luo et al. [86]propose a medical image zero-watermarking algorithm that utilizes the histogram of oriented gradients (HOG) and discrete cosine transform (DCT). The algorithm first extracts the feature vector of the medical image and then generates a corresponding feature sequence using the perceptual hash algorithm. The watermark image is subsequently encrypted using chaotic scrambling through logical mapping. Fang et al.[87] propose a novel zero-watermarking algorithm for medical images that utilizes the bandelet transform and discrete cosine transform (Bandelet-DCT). As a preprocessing step, scale-invariant feature transform (SIFT) is applied to the original medical image to extract features. The watermark, containing patient information, is then encrypted using the chaotic tent map. Bandelet-DCT is subsequently used to extract visual feature vectors from the medical images. Finally, watermark embedding and extraction are achieved by combining zero-watermarking technology with cryptography.

Sun et al. [88] propose a robust zero-watermarking algorithm for medical images that utilizes the AGAST-LATCH feature extractor and discrete cosine transform (DCT). The AGAST algorithm and LATCH algorithm are combined to extract the feature matrix of the medical image. DCT and perceptual hashing are then applied to the feature matrix to obtain hash sequences. These hash sequences are combined with encrypted watermark information, which is secured using the chaotic state of the Logistic Map.

Seenivasagam and Velumani [89] propose a zero-watermarking scheme for unambiguous medical image authentication based on the composite contourlet transform (CT) and singular value decomposition (SVD). A framework is also presented for accessing patient records using this watermarking scheme. The watermark consists of patient identification details and a link to patient data encoded in a QR code. The proposed scheme ensures that the medical image remains unaltered by the watermarking process. Patient authentication and authorized access to patient data are achieved by combining a secret share with the master share, which is constructed from invariant features of the medical image. Hu's invariant image moments are used to create the master share. Liu et al. [19] propose a novel robust zero-watermarking algorithm for medical images that utilizes scale-invariant feature transform (SIFT) and discrete cosine transform (DCT) in the encrypted domain. The original medical image is first encrypted in the transform domain using a Logistic chaotic sequence to enhance concealment. SIFT-DCT is then applied to the encrypted medical image to extract feature sequences. Finally, a zero-watermarking technique is employed to ensure that the region of interest (ROI) of the medical image remains unchanged.

Magdy et al. [90] propose fast multiple zero-watermarking methods for medical image security and copyright protection in Internet of Medical Things (IoMT) applications. These methods utilize multi-channel fractional Legendre Fourier moments (MFrLFMs) to ensure the original medical images remain undeformed. MFrLFMs are chosen for their high accuracy, numerical stability, geometric invariance, and resistance to attacks. The most significant features extracted from MFrLFMs are scrambled using the two-dimensional Discrete Henon Map and then XORed with a binary scrambled watermark to create an owner share. The proposed watermarking method is implemented on a low-cost Raspberry Pi Linux microprocessor, making it suitable for medical devices in IoMT environments.

Lastly, Fang et al. [44] proposes a novel fragile zero-watermarking algorithm that utilizes quaternions, local binary patterns (LBP) and SVD. This combined approach, leveraging both blockchain and fragile zero-watermarking, can potentially address the limitations of cross-chain prosecution and confirmation protection. Collectively, these contributions demonstrate a significant advancement in the field of medical image watermarking, providing robust solutions to safeguard patient data against unauthorized access and potential tampering. Another method proposes a zero-robust digital watermarking algorithm for medical images, based on the combination of Zernike moments and discrete cosine transform (DCT). Yang et al. [18] begins by extracting key features from the medical image using Canny edge detection, followed by processing the contour points with Zernike moments to obtain a robust representation invariant to common geometric transformations.

Then, the DCT is applied to decompose the image into frequencies and a 32-bit visual feature vector is extracted. The digital watermark, previously encrypted using a chaotic sequence, is then combined with this vector to generate a unique key stored in a secure location. During verification, the suspected medical image undergoes the same feature extraction process, and the original key is used to retrieve and decrypt the digital watermark, thus confirming the authenticity of the image.

Here's a summary of the performance of the mentioned zero-watermarking methods in this table:

Table 3 . Performance of mentioned zero-watermarking methods

Method	**Technique**	**Image Modalities**	**Metrics**	**MI +W**	**Pros**	**Cons**
Fang et al. [87]	Bandelet-DCT, SIFT, chaotic tent map	Medical images	NC	512×512 + 32×32	Robust against geometric attacks, high security	Requires complex feature extraction
Seenivasagam and Velumani [89]	Composite contourlet transform (CT), SVD	CT, Mammogram, MRA, PET,Ultrasound, Nuclear, and X-ray	NC, BER, SSIM, UIQI	512×512 + 77×77	Versatile for various medical image types, provides patient authentication	Computationally intensive
Liu et al. [19]	SIFT, DCT, Logistic chaotic sequence	abdominal CT image	NC, PSNR	/ + 32×32	Robust against attacks, efficient	have higher time complexity with chaotic sequence and SIFT processing.
Magdy et al. [90]	MFrLFMs, Discrete Henon Map	Medical images	NC, BER, SSIM,PSNR	256×256 + 32×32	Fast and suitable for IoMT devices, high accuracy	require specialized hardware for MFrLFMs
F.liu et al. [44]	Quaternions, LBP, SVD, blockchain	/	/	513×513 + 57×57	High security with blockchain integration. Robust against tampering.	Complex implementation, require significant computational resources
Arevalo-Ancona and Cedillo-Hernández [85]	K-means clustering, DFT	Medical images	BER = 0.0058, NC= 0.9944	515×515+160×160	Efficient ROI detection, robust against geometric attacks	sensitive to image noise and to the initial selection of cluster centroids.
Yang et al. [18]	Zernike, DCT	Medical images	PSNR, NC	/ + 32×32	Simple implementation, robust against common geometric transformations.	High complexity due to Zernike moments.

III.4. Deep Learning-Based Watermarking Techniques

The process of embedding and extracting watermarks leverages the power of neural networks to learn complex features from images, providing more sophisticated and often more robust watermarking solutions compared to handcrafted techniques. Deep neural networks have revolutionized image watermarking, providing new avenues for robust and efficient techniques. Our primary focus in this thesis is the use of deep-learned features, which have demonstrated remarkable potential in improving the security, imperceptibility, and robustness of medical image watermarking. To better understand these techniques, we categorize them into three primary groups based on their application in recent research: Joint Training, Feature Transformation and Hybrid.

III.4.1. Joint Training Methods

As illustrated in Fig. 20, joint training methods involve training an embedder network to integrate a watermark into a cover image and an extractor network to retrieve the embedded watermark. Some variations incorporate separate feature extraction networks within the embedder for pre-processing.

To enhance robustness, a noise module is often included after the marked image. This module introduces noise during training, helping the extractor network become more resilient to disturbances. All components are typically trained jointly in a deep neural network framework using gradient descent.

The goal is to minimize the variance between the original and extracted watermarks while maintaining imperceptibility. The loss function often consists of two terms:

- Image fidelity: Measures the visual difference between the original and marked images.
- Watermark extraction accuracy**:** Evaluates how well the extractor can recover the embedded watermark.

Balancing these two terms is crucial. Excessive focus on image fidelity can limit the available space for watermark embedding, hindering extraction. A common approach involves using gradients from the watermark extraction term to refine the weights of all

components, while gradients from the image fidelity term are used only to optimize the embedder network.

The following sections will delve into specific joint training methods and their unique characteristics.

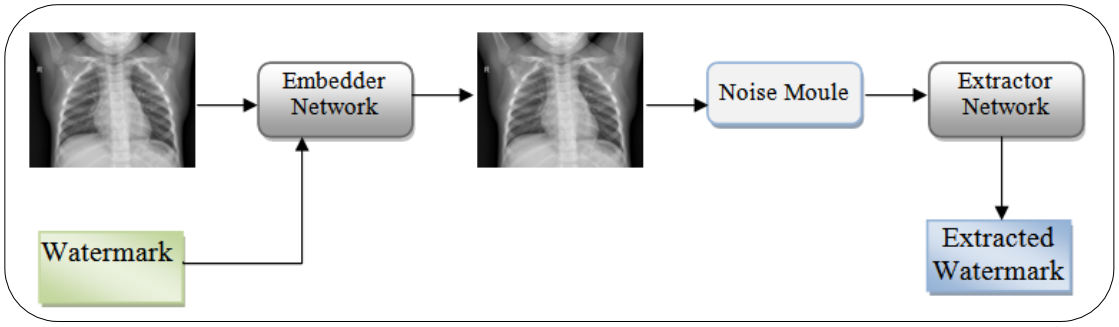


Fig 20. General process of the embedder–extraction joint training.

The concept of joint training, introduced in HiDDeN (hiding data with deep networks) by Zhu et al. [91], revolutionized image watermarking by integrating the embedder-extractor paradigm. This innovative approach streamlined the processes of image steganography and watermarking, paving the way for more efficient and robust techniques.

HiDDeN incorporated several key design elements:

- Watermark Replication: The embedder replicated the watermark to ensure consistent embedding.
- Final Layer Embedding: Watermark integration at the final layer preserved visual quality.
- Discriminator Networks: These networks ensured the embedder generated visually coherent marked images.
- Noise Layer: The noise layer introduced various noise types to enhance robustness against attacks.

To address the non-differentiability of JPEG compression, HiDDeN introduced a differentiable JPEG variant, enabling seamless gradient flow during training.

Singh et al. [92] propose a novel CNN-based watermarking technique for digital images. The encoder network extracts latent features from the cover and secret images, which are subsequently concatenated to generate the watermarked image. A denoising autoencoder is employed by the recipient to remove noise from the watermarked image. Finally, a CNN is

used to extract the secret watermark.Zhang et al. [93] proposed Universal Deep Hiding (UDH), another pioneering work in joint training. UDH was among the first to embed entire images as watermarks within embedder-decoder frameworks. Its novel encoding approach facilitated disentanglement during extraction, ensuring the integrity of the cover image.

Sinhal et al. [94] propose a robust, blind watermarking technique for medical images utilizing a deep neural network (DNN). To preserve the region of interest (ROI) data in the original image, the LZW compression method is employed during watermarking. The IWT (Integer Wavelet Transform) is used to embed the watermark. The SHA-256 technique generates hash keys for the ROI and RONI (region of non-interest) areas. A deep neural network-based architecture enables robust watermark extraction. Jaiswal et al.[95] introduced a DNN-based approach for blind watermarking of color images. By applying integer wavelet transform (IWT) and principal component analysis (PCA), they transformed and reduced the input image's features. Watermark images were encoded in the blue color channel, while a binary classification technique was used for watermark extraction. The DNN was trained using a 512 × 11 training set and a 512 × 10 testing set.

Rai et al. [96] introduced an improved chimp optimization (ECO) technique designed for robust watermarking, leveraging the power of deep fusion convolutional neural networks (DFCNN). This approach employs a dual-network architecture, consisting of an embedding network and an extraction network. The embedding network, utilizing octave convolutional layers, effectively reduces spatial redundancy while capturing essential features. In contrast, the extraction network's pyramid feature extraction module focuses on extracting local features, and the use of dilated convolutions helps to minimize model parameters.

A significant challenge in joint training is the need for a differentiable noise layer. Liu et al. [97] addressed this with a two-stage training method:

1. Collaborative Training**:** Both embedder and extractor were trained without noise to establish a strong foundation.
2. Extractor-Only Training: The embedder's parameters were fixed, allowing for the introduction of non-differentiable noise in the extractor.

. For momentum-based optimization, joint training is not strictly necessary. However, the embedder should generate high-quality images, and the extractor should be capable of

retrieving features post-JPEG noise. Zhang et al. [98] introduced a pseudo-differentiable method for JPEG compression to address these requirements. Chen et al. [99] used simulation networks to emulate JPEG compression, while Jia et al. [60] incorporated both actual and simulated JPEG compressions in batches.

To address specific challenges, such as camera resampling noise, Fang et al. [100] and Gu et al. [101] introduced screen-shooting noise layer simulations. These simulations helped models become more robust to real-world distortions. The Huang et al. technique [102] is an innovative approach for embedding watermarks into digital images using Generative Adversarial Networks (GANs). This method relies on several key components: Feature Fusion Module (FFM) for extracting multi-layer and deep features from the image, an Attention Module (AM) to identify suitable regions for embedding the watermark, and a discriminator to train the network to differentiate between original images and those with embedded watermarks. In this technique, the discriminator's input is a pair of images: the original image and the watermarked image. The goal of the discriminator is to learn to distinguish the subtle differences between the two images, thereby verifying the presence of the watermark in a new image.

Existing joint training methods often require explicit identification of training noise. Zhong et al. [21] introduced an invariance layer to filter out extraneous information, making models more robust without relying on predefined noise lists. Adversarial networks can also enhance robustness. Luo et al. [103] incorporated an adversarial network as a noise module, forcing the extractor to counter adversarial perturbations.

The joint embedder-extractor paradigm has been a highly effective approach. Ahmadi et al. [104] introduced a variety of noise types in their noise layer. Plata et al. [105] proposed a propagator to disseminate the watermark across the image. Zhang et al. [106]introduced a multi-class discriminator. Hao et al. [107] proposed a high-pass filter at the discriminator's inception. Xu et al. [108] used a reversible neural network. Mahapatra et al. [109] incorporated the difference between the marked and cover images into the extractor. Zhao et al. [110] introduced a factor to modulate the watermark's intensity and used a spatial attention feature map. Ying et al. [111] targeted embedding capacity and used a decoupling and revealing network. Fang et al. [112] introduced an extractor-embedder-extractor architecture. Mun et al. [113] incorporated reinforcement learning. Zhu et al. [114] propose MDResNet-HDWM, a blind watermarking technique for high-definition images. This method leverages

deep learning and key-point detection to securely embed watermarks in multiple non-overlapping regions. By normalizing the image and embedding watermarks centrally within these regions, we ensure scale-invariant protection. MDResNet, our trained neural network, uses a curriculum learning approach to enhance robustness against various image manipulations, including signal processing and geometric transformations.

Multimedia watermarking has also been explored. Das and Zhong [115] embedded audio watermarks into cover images. Ge et al. [116] proposed a technique for document images with multiple skip connections. Liao et al. [117] extended watermarking to GIF animations using 3D deep neural networks. Since the introduction of HiDDeN, the field of joint training has seen significant advancements. Researchers have addressed various challenges and proposed innovative solutions, making deep learning-based image watermarking a dynamic and evolving field.

Table 4 provides a comprehensive summary of these challenges and representative solutions.

Table 4. Challenges and Representative Solutions in Embedder-Extractor Image Watermarking [8]

Challenge	**Representative Solution**
Differentiable Noise Layer	Two-stage training scheme, differentiable JPEG simulations
Non-differentiable Nature and Low-Performance Issues of JPEG	Differentiable JPEG simulations
Special Challenging Noises (e.g., camera resampling)	Simulated camera distortions in the noise layer
Lack of Robustness to Untrained Noises	Invariance layer, adversarial networks
Enhanced Overall Model Performance	Innovative architectures, training paradigms, feature transformation
Multimedia Cover Images	Specialized neural networks for multi-modal features, robustness

Gao et al. [118] propose a novel framework for robust image watermarking based on a generative adversarial network (RIW-GAN). In their approach, the encoder network consists of convolutional layers and a residual block, which produces an encoded image with minimal distortion, closely resembling the original image. To improve the model's resilience to attacks, a simulated noise layer is incorporated as a differentiable network layer, facilitating end-to-end training prior to decoding.

III.4.2. Feature Transformation:

As illustrated in Figure 3, this category of watermarking methods primarily utilizes deep neural networks for feature transformation. Both cover and marked images undergo transformations within these networks, creating distinct feature spaces. Watermark embedding and extraction are then performed within these spaces.

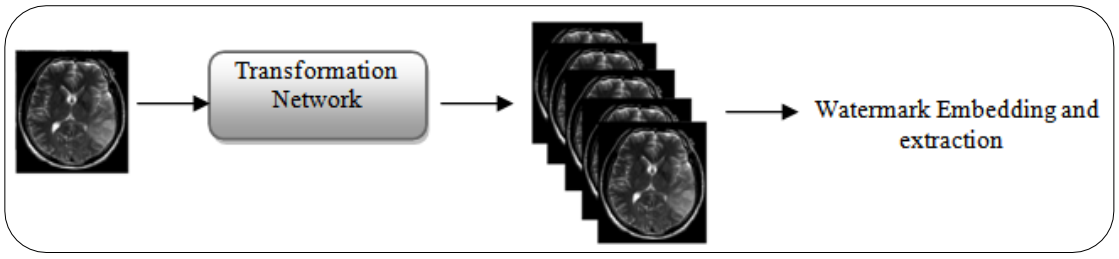


Fig 21. General process of the deep networks for feature transformation.

A key expectation is the robustness of the transformed domains. Minor alterations to the marked images should yield consistent or nearly identical feature values.

Numerous methods have adopted deep networks for feature transformation in deep learning-based zero watermarking. Fierro et al. [119] used CNNs to extract features from cover images, which were then integrated with a permuted binary watermark sequence via an XOR operation to create a master share. The same CNN processes a test image to extract features, which are subsequently XORed with the master share to extract the watermark. An appropriate key ensures the identification of the watermark.

Huang et al.'s watermarking technique [62] employs a depth wise over parameterized VGG (DO-VGG) model and an attention mechanism to embed watermarks into medical images without compromising image quality. The DO-VGG extracts deep, high-dimensional features from the image, which are then enhanced using a channel and spatial attention module (CBAM) to focus on the most informative regions. The watermark is encrypted using an enhanced logistic chaotic map and embedded into these refined features. This approach ensures that the watermark is robust against various attacks while preserving the original image quality.

He et al. [120] extended this approach by adding fully connected layers and introducing a shrinkage module for feature extraction precision. They also focused on eliminating redundancy in the feature space and incorporated a noise layer during feature

training to increase robustness. Han et al. [121] enhanced this methodology by introducing a chaotic encryption algorithm and using the Swin Transformer for feature generation. The Zhang et al. algorithm [22] proposes a robust watermarking technique for medical images using GoogLeNet transfer learning. A pre-trained GoogLeNet network is fine-tuned on the medical dataset to extract stable feature vectors. Multiple types of watermark information are encrypted using two-dimensional Henon chaos encryption. The encrypted watermarks are combined with the feature vector to create a key, ensuring zero-watermarking and protecting patient data privacy.

Another research direction involves employing pre-trained deep neural networks. The input data is trained to yield the desired marked images while the pre-trained weights remain static. The resulting marked image is visually similar to the original cover image but reveals the embedded watermark upon feature extraction. Vukotic et al. [122] illustrated this by implementing a pre-trained CNN and adaptively modifying the input cover image through gradient descent. The Haimour et al. Study [123] applied transfer learning to translate MR-CT images and vice versa using 18 pre-trained non-medical models. The models were fine-tuned to achieve the best results. Fernandez et al. [124] extended this approach to multi-bit extraction by assigning distinct keys to each bit of the binary watermark sequence. They used self-supervised learning for the pre-trained networks, which offers advantages in terms of feature space robustness.

The use of deep networks for feature transformation is a relatively new approach in image watermarking. It differs from the more conventional joint training model, which aligns more closely with traditional image watermarking paradigms. As a result, the academic literature on this method is limited. However, it offers promising challenges**:**

1. Robustness of Transformed Features**:** Ensuring that the transformed features remain robust to various attacks and distortions is crucial. Techniques such as noise layers, feature extraction techniques, and careful feature space design can help address this challenge.
2. Efficiency of Feature Extraction**:** The efficiency of the feature extraction process is essential, especially for real-time applications. Employing efficient deep network architectures and optimization techniques can improve efficiency.

3. Security of Watermark Embedding: Protecting the embedded watermark from unauthorized extraction or tampering is a critical concern. Chaotic encryption, feature space design, and other security measures can be implemented.
4. Compatibility with Different Watermark Types: The method should be compatible with various watermark types, including binary, grayscale, or color images. Adaptable feature extraction techniques can address this challenge.
5. Generalizability to Various Image Domains: The method should be generalizable to different types of images, such as medical images, natural images, or document images. Diverse training data and transfer learning can enhance generalizability.

By addressing these challenges, researchers can develop more effective and robust feature transformation methods for image watermarking.

III.4.3. Hybrid Methods:

Hybrid methods, which combine deep learning techniques with traditional image watermarking calculations, offer a powerful approach to enhancing the performance and robustness of watermarking systems. By leveraging the complementary strengths of both domains, these methods can achieve significant improvements over traditional techniques alone.

Chinnamuthu et al. [125] present a novel technique in digital watermarking that incorporates data hiding through segmentation and classification using deep learning methods. The input images are medical images, including Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans, which undergo processing for noise removal, smoothing, and normalization. The processed images are then watermarked using a Singular Value Decomposition-based discrete wavelet transform quantization model, with segmentation and classification achieved through convolutional generative adversarial networks. Kandi et al. [126] employed two convolutional autoencoders to reconstruct a cover image individually. The distinctions between the reconstructed images and the original cover image were used to represent bits in a binary watermark. Ferdowsi et al. [127] introduced a technique for IoT applications using classic spread spectrum for watermark embedding and mapping cover image features to bit streams.

Singh et al. [128] introduce a GAN-based image watermarking application tailored for healthcare scenarios**.** The medical image undergoes initial encryption using randomized singular value decomposition (RSVD) and a 3D chaotic map. Subsequently, a GAN is employed to create a final watermark by embedding the patient's identifier and the image hash into the hospital logo. The encrypted medical image is then watermarked using RSVD-based embedding and redundant discrete wavelet transform (RDWT). Experimental results demonstrate the undetectability and robustness of the proposed approach against various attacks. While the design paradigms and operational frameworks of hybrid methods can vary, a common approach involves using deep learning for watermark extraction. Deep learning models, with their ability to fit complex functions and uncover intricate patterns, are well-suited for the challenging task of extracting embedded watermarks from watermarked images. Traditional image watermarking calculations, on the other hand, provide a solid foundation and offer stability, reliability, and interpretability.

Li et al. [61] integrated pre-processed grayscale watermark images into the DCT blocks of cover images and used a CNN for extraction. Mellimi et al. [129] embedded watermarks into the lifting wavelet domain and used a deep neural network extractor. Chack et al. [130] introduced a hybrid methodology combining traditional watermarking, CNN, and evolutionary optimization. Fang et al. [131] presented a deep template-based image watermarking mechanism. Kim et al. [132] introduced a template-based approach with patch segmentation and curvelet domain watermarking. Chen et al. [133] developed a mechanism for authenticating watermark systems using deep learning. Mahto et al. [134] [73] developed a robust and imperceptible watermarking technique based on hybrid optimization. Their approach combines spatial and transform domains to embed hidden marks within the cover material. By employing a multi-type mark with appropriate scaling, they enhance the algorithm's legitimacy. To further strengthen security, the watermarked image is encrypted using an advanced encryption method. Finally, a Denoising CNN is integrated to improve the algorithm's resilience against noise and attacks.

Hybrid methods often encounter challenges due to the complex interplay between deep learning and traditional watermarking. These challenges and their representative solutions are summarized below:

- Determining the optimal role of deep learning: Deep learning can be employed for various tasks, such as watermark extraction, feature enhancement, or other auxiliary

functions. Careful analysis is required to determine the most effective role for deep learning in a given hybrid method.

- Increased complexity: The integration of deep learning and traditional techniques can introduce additional complexity into the system. Modular and scalable architectures can help manage this complexity and facilitate integration.
- Balancing deep learning and traditional techniques: Striking a balance between the strengths of deep learning and traditional techniques is crucial for optimal performance. Careful consideration of the complementary nature of these approaches is essential.
- Integration with existing watermarking systems: Hybrid methods must be compatible with existing watermarking systems. This may require adaptation strategies or modifications to the existing system.
- Evaluation and benchmarking: Comprehensive evaluation metrics and comparison with state-of-the-art methods are necessary to assess the effectiveness of hybrid methods.

By addressing these challenges and leveraging the strengths of both deep learning and traditional watermarking techniques, researchers can develop highly effective hybrid methods that advance the state-of-the-art in image watermarking.

Table 5.Summary of some mentienned methods:

Truth	**Deep Role**	**Purpose**	**Strategies Employed**	**Domain**	**Performance Assessment**
(Chacko and Chacko. [130]	Watermark identification and extraction	Optimized positioning and parameters	DCT, Arnold transform, HHO, DLCNN	Hybrid	NC, BER (attacks), PSNR= 46 dB
W. Zhang et al. [22]	Feature extraction	Resilient medical watermarking algorithm	GoogLeNet, Chaotic Henon mapping	Feature Transformation	NC, PSNR (attacks)
Zhu et al. [114]	Watermark embedding and extraction	High-definition watermarking scheme	MDResNet, Embedding-revealing network	Joint Training	NC=0.9685, BER=0.0343, PSNR=38.8790, SSIM=0.9581
Singh and Singh. [92]	Watermark embedding and extraction	Minimize human involvement	Encoder-decoder network : CNN	Joint Training	NC=0.9996, PSNR=44.48 dB, SSIM=0.9997
Jaiswal et al.	Watermark extraction.	Enhanced imperceptibility-	IWT, PCA, DNN	Joint Training	Average NC= 0.9953, Average PSNR= 35.57 dB, SSIM=0.9969

[95]		robustness trade-off			
Mahto et al. [134]	Denoising recovered watermark image	Improved watermarking through hybrid optimization	LWT-Schur-T-SVD, HPSOF, SIE, CNN	Hybrid	NC and BER approaching 1 and 0, respectively, Average PSNR=57.7124 dB
Sinhal and Sinhal et al. [94]	Watermark extraction.	Offer strong security measures for medical images.	LZW, IWT, SHA-256, DNN	Joint Training	NC=1, BER=0, PSNR=40.27, SSIM=0.9445
T. Huang et al. [62]	Feature extraction	Lossless medical watermarking resistant to geometric attacks	VGG (DO-VGG), CBAM	Feature Transformation	NC, PSNR (Attacks)

III.5. Zero Watermarking Based on Deep Learning

Zero watermarking techniques offer an innovative approach by embedding watermarks without directly modifying the host image. In the context of deep learning, this process leverages feature extraction from images, using neural networks to learn complex representations that can securely embed and recover watermarks. Since the original image is not altered, zero watermarking preserves its visual integrity while ensuring the robustness of the watermark against attacks, similar handcrafted methods.

The following methods employ deep learning for feature extraction (FE), for instance Liu et al.'s watermarking method [135] is a novel approach that embeds watermarks into images without significantly altering their original appearance. Instead of directly modifying the image pixels, it combines the watermark information with the style of the host image. A convolutional neural network (CNN) is used to transfer the style of the host image onto the watermark, creating a stylized image. The resulting image is then encrypted using the Arnold transform and a timestamp is added for additional security. To verify the watermark's presence, another CNN is trained to distinguish between the original watermark and the stylized image. Darwish et al. [136] propose a zero-watermarking method for color images utilizing CNN and 2D logistic-adjusted Chebyshev map (2D-LACM). Pre-trained VGG19 extracts deep feature maps from the original image. These feature maps are combined to form a feature image, and the owner's watermark sequence is XORed with this image. 2D-LACM encrypts the watermark and scrambles the feature matrix for enhanced security. Experimental results demonstrate the proposed approach's invisibility and resilience.

Nawaz et al.'s zero-watermarking technique [26] utilizes DWT and DCT transforms, combined with a ResNet deep neural network, to embed a watermark into medical images without visible alterations. The method extracts deep features from the image using ResNet, which are then transformed using DWT and DCT. A perceptual hash function is applied to these transformed features to create a feature vector. The watermark is embedded within this feature vector, and the resulting vector is encrypted. Xiang et al.'s zero-watermarking technique [137] leverages deep neural networks, specifically residual networks, to embed a watermark into medical images without visible alterations. The method extracts high-level features from the image, which are then used to create a style-based watermark. Rather than using the traditional XOR operation for watermark verification, a residual network is trained to iteratively extract the watermark from the zero-watermarked image, improving its quality with a defined loss function.

Gong et al. [23] propose a zero-watermarking approach using Residual-DenseNet to embed a watermark into medical images without visible alterations. The method employs a CNN to extract resilient feature vectors from the image. These features are then combined with a watermark that is encrypted using a chaotic matrix generated by the logistic map. The resulting watermark is embedded into the image without modifying the original pixel values. This approach demonstrates strong robustness against both conventional and geometric attacks, as confirmed by experimental results. In addition, Nawaz et al. .[25] introduces again a novel zero-watermarking technique that combines DWT, an improved MobileNetV2 CNN, and DCT to enhance the robustness of watermarking in encrypted medical images. The method involves extracting features from encrypted images using a modified MobileNetV2, transforming these features with DWT and DCT, and embedding a watermark encrypted using a logistic map and hash function.

Li et al.[138] proposes a novel zero-watermarking technique that leverages DCT and a pre-trained DarkNet53 network to enhance robustness in encrypted medical images. The method involves extracting 32-bit features from the encrypted image using DCT, and 128-bit features from the same image using a modified DarkNet53. These features are then combined to create a robust feature vector, which is used for watermark encryption. Anand et al. [139] proposes a novel zero-watermarking technique for securing healthcare records. The method involves visibly marking the carrier image with a hospital logo, which is then scrambled using a space-filling curve. AlexNet is employed to extract deep features from the visibly marked image. This

approach ensures that the watermark is robust against various attacks while providing a visible indication of ownership.

Table 6. illustration of zero watermarking methods

Authors	Deep Role	Strategies Employed	Domain	Performance Assessment
Liu et al. [135]	FE, Watermark detection	Zero-watermark image construction, Arnold transform, CNN	Joint Training	correlation coefficient (attacks), Correlation coefficients for ten watermark logos
Darwish et al. [136]	FE	VGG19, 2D-LACM	Feature Transformation	BER below 0.0044, NC above 0.9929, Average PSNR=33.1537
Nawaz et al. [26]	FE	ResNet ,DWT, DCT, chaotic scrambling	Hybrid	PSNR (attacks), Most NC larger than 0.50
Xiang et al. [137]	FE, Replace traditional XOR operation	ResNet : Generator network and detector network	Joint Training	NC, PSNR (attacks)
Gong et al. [23]	FE	Residual-DenseNet	Feature Transformation	NC, PSNR (attacks)
Dong et al. [140]	FE	NasNet-Mobile, DCT, chaos map, Arnold transform	Hybrid	NC, PSNR (attacks)
Ref.[25]]	Processing the watermark of images	MobileNetV2,DWT, DCT, logistic map	Hybrid	NC (attacks)
Li et al. [138]	FE	DarkNet53, DCT, Tent mapping	Hybrid	NC, PSNR (attacks)
Anand et al. [139]	FE	NSST, SVD, SSFC, Alexnet	Hybrid	NC, PSNR (attacks)

III.6. Conclusion

This chapter has presented a comprehensive review of the state of the art in image watermarking, focusing on the comparison between handcrafted and deep learning-based methods. We have explored the principles and techniques underlying both approaches, analyzed their advantages and limitations, and discussed their applications in the context of medical image security.

Handcrafted watermarking techniques have been widely used for many years, and they offer certain advantages such as interpretability and computational efficiency. However, deep learning-based methods have demonstrated significant potential in recent years, with their ability to learn complex patterns and features from data.

By carefully considering the specific requirements of a medical image watermarking application, researchers and practitioners can choose the most appropriate method. In some

cases, a hybrid approach combining handcrafted and deep learning techniques may be beneficial. As research in this field continues to advance, we can expect to see further innovations and improvements in image watermarking techniques for medical image security.

IV. Chapter 3 : Proposed Methods

IV.1. Introduction

In this chapter, we propose three novel methods for medical image watermarking, focusing on enhancing feature extraction techniques to address the challenges of security, robustness, and imperceptibility in telemedicine applications. With the growing need to protect sensitive medical data, watermarking techniques must be robust against image processing attacks while maintaining high image quality. By improving feature extraction in the watermarking process, the proposed methods aim to embed and retrieve watermarks more effectively across different conditions.

This chapter proposes three medical image watermarking methods. The first is a blind watermarking approach using LBP with DWT to capture local texture patterns, enhancing robustness against noise and filtering. The second method incorporates gradient analysis into the DWT framework, making the watermark more resilient to attacks like blurring and compression. The third method uses deep learning with a pre-trained CNN for zero watermarking, embedding the watermark in the image's feature space without altering the original image. Each method is evaluated using PSNR, NC, and tested against common attacks.

IV.2. Medical Image Watermarking Based on LBP-DWT

we propose a medical image watermarking method that combines LBP with DWT, LBP is known for its ability to capture local texture information, making it suitable for extracting key features from medical images. Meanwhile, DWT provides a multi-resolution analysis, allowing us to embed the watermark in a way that is resilient to common attacks such as noise, compression, and geometric distortions. The combination of these two methods allows for the development of a robust watermarking system tailored to the unique requirements of medical images.

IV.2.1. Background Concepts

To understand this approach in detail, it is important to review the underlying background concepts that contribute to its effectiveness:

a) Local Binary Patterns (LBP)

➢ *Definition*:

Local Binary Patterns (LBP) is a texture descriptor used in image processing and computer vision to describe local features by comparing the pixel values around a central pixel in a defined neighborhood[49].

➢ *How It Works:*

- For each pixel in the image, a 3x3 block (or larger) is considered.
- The center pixel is compared with its surrounding neighbors. If a neighboring pixel value is greater than or equal to the center pixel, the value 1 is assigned; otherwise, the value 0 is assigned.
- These binary values are combined to form an 8-bit binary number (LBP code) for each pixel, which is then converted into a decimal value.

Example: For a 3x3 block as shown in Fig. 21, the center pixel is compared with the 8 neighboring pixels, resulting in a binary pattern. This pattern is used for feature extraction, especially in tasks like texture classification.

$$LBP_j = \sum_{i=0}^{i=7} S(P_c, P_i) * 2^i \tag{8}$$

Where P_c represents the value of the central pixel in block j, and Pi represents the corresponding pixel values. The sign function, denoted as S and defined as (Eq. 9), is expressed as:

$$S(P_c, P_i) = \begin{cases} 1, & if \ \ P_c \geq P_i \\ 0, & else \end{cases} \qquad \text{Where: } i \in \{1, \ldots, 8\} \tag{9}$$

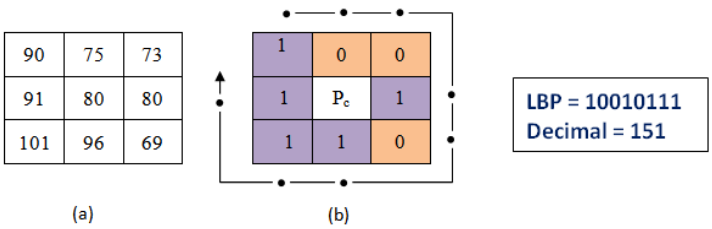


Fig 22. Illustration of LBP mechanism

b) Zigzag Scanning

➢ *Definition:*

Zigzag scanning is a method used to rearrange the elements of a matrix or block in a diagonal sequence. It is commonly used in image compression algorithms like JPEG to convert a 2D matrix into a 1D vector[47].

➢ *How It Works:*

- The elements of a matrix are traversed in a Zigzag pattern starting from the top-left corner, moving diagonally down and to the right until the entire matrix is covered.
- This method is particularly useful for reordering high-frequency and low-frequency components in an image block, which helps in efficient compression and embedding.

Fig. 23 illustrates the Zigzag process used to prepare the watermark segments.

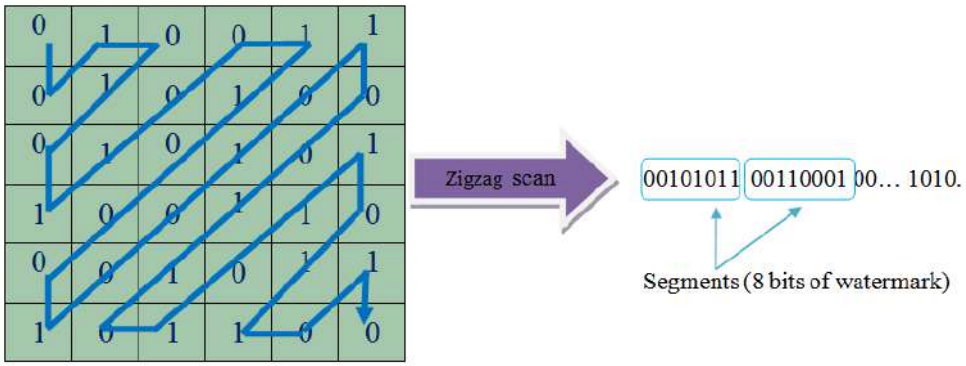


Fig 23. Illustration of Zigzag scan and segmentation of Watermark

c) Arnold Transform

➢ *Definition:*

The Arnold transform is a scrambling algorithm used to distort and encrypt an image by rearranging the pixel positions in a specific pattern. It is often applied to watermarks to enhance security[141].

➢ *How It Works:*

- The transform repeatedly scrambles the pixel positions based on a mathematical formula, making the image appear disordered. The process can be reversed using the inverse Arnold transform if the key (number of iterations) is known.

Formula: For a 2D image of size N×N, the Arnold transform is defined by:

$$\begin{pmatrix} i' \\ j' \end{pmatrix} = \begin{pmatrix} 1 & a \\ b & ab+1 \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix} (mod\ N) \tag{10}$$

Where: (i, j) represents the original pixel positions in the image, (i', j') are the positions that correspond to the following permutation. The control parameters a and b allow for pixel position adjustments. For different image sizes and parameters, The period T in the Arnold transform varies for different image sizes and parameter settings, where T depends on the image size [141],and it determines the number of permutations required to return the image pixels to their original positions. The use of the Arnold transform not only scrambles the image but also provides enhanced security through the use of the T value as a key in our scheme[28-29]

IV.2.2. Generation of Watermark

The patient's information is converted into a binary image, which serves as a text-based watermark for embedding into the medical image, as shown in Fig. 24.

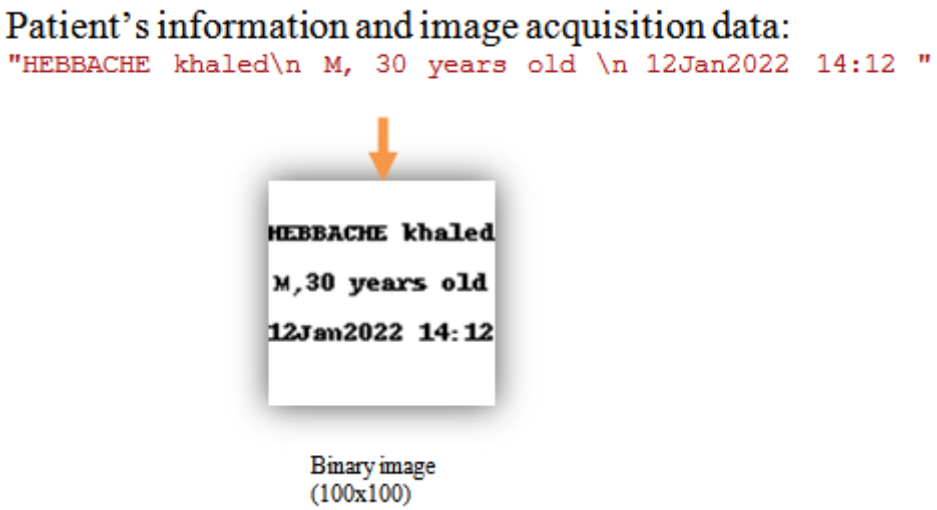


Fig 24. Creation of watermark

IV.2.3. Embedding Process

The embedding process involves several key steps:

1. Normalization and Transformation: The medical image is normalized and transformed into the frequency domain using DWT to obtain the sub-bands (LL, LH, HL, HH).
2. Selection of Sub-band: The suitable sub-band for embedding is identified as LL.
3. Partitioning and Preparation:
    - The LL sub-band is divided into non-overlapping 3x3 blocks. Each block is multiplied by 10, with the reference coefficient for each block denoted as P.
4. Extraction of LBP code:
    - For each block P, eight bits of LBP code are extracted using the following steps:
        - Split Block: The block P is divided into two parts: PI (integer values) and PF (fractional values).
        - LBP Code Extraction: The LBP code is extracted from PI using specified equations (1) and (2).
5. Watermark Transformation:
    - The watermark W is scrambled using the Arnold transform with a secret key, resulting in the scrambled watermark W^A.
6. Vector Conversion:
    - The scrambled watermark W^A is converted into a 1D vector using the Zigzag scan method, then divided into 8-bit segments.
7. Encryption with LBP:
    - Each segment of the watermark W_i^A is encrypted with the corresponding LBP code using the XOR operation:

$$X_j = W_i^A \oplus LBP_j \tag{11}$$

8. Bit Pair Switching:
    - The bits in X_j are organized into pairs, with each pair's bits switched, resulting in the code Y_j:

$$Y_j = \{ y_j^0 = x_j^1; y_j^1 = x_j^0; y_j^2 = x_j^3; y_j^3 = x_j^2; y_j^4 = x_j^5; y_j^5 = x_j^4; y_j^6 = x_j^7; y_j^7 = x_j^6 \} \tag{12}$$

9. Embedding:
   - The bits from Y_j are embedded into the corresponding bits of PI using the Least Significant Bit (LSB) method, yielding a watermarked block PI'.
   - To maintain local neighborhood relationships, the coefficients in PI' are adjusted using the following modification:

$$PI_i^* = PreserveLBP\,(PI_c, PI_i, PI_i') = \begin{cases} PI_i' + 2, & if\ (PI_c \geq PI_i)\ AND\ (PI_c < PI_i') \\ PI_i' - 2, & if\ (PI_c < PI_i)\ AND\ (PI_c \geq PI_i') \\ PI_i' & otherwise \end{cases} \quad (13)$$

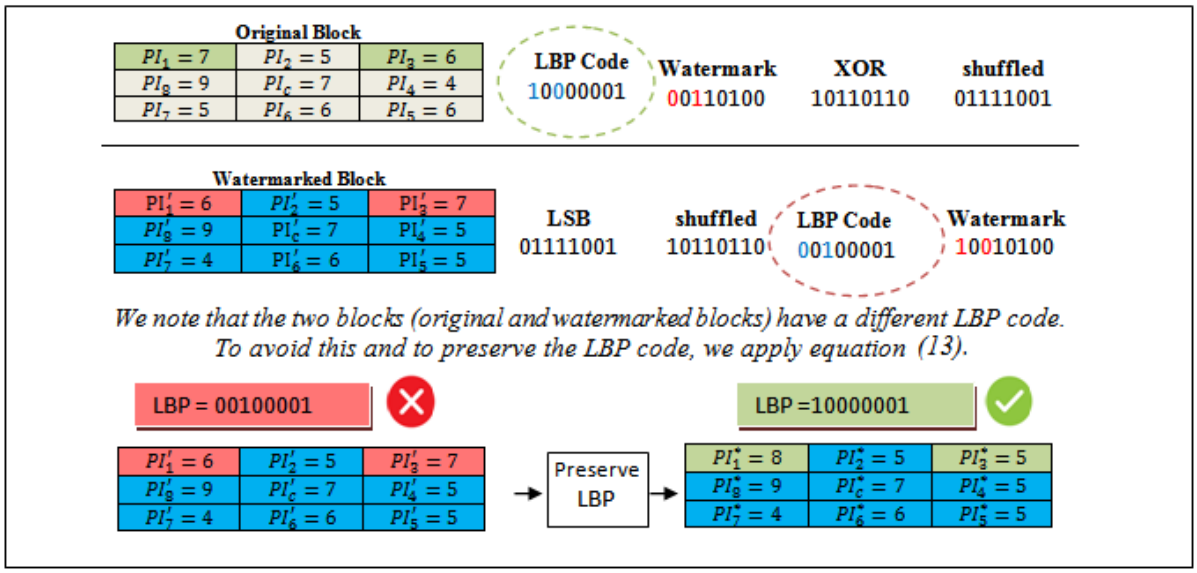


Fig 25. the preservation process of Block (Eq. 13)

However, this process can introduce slight changes to the block's pixel values, which could potentially alter the original LBP code. Since LBP extraction relies on these pixel values, any such change would lead to a mismatch between the originally embedded code and the one extracted later, even without any external attack.

Our analysis revealed changes (Fig. 25) in specific pixel intensities values ($PI_1, PI_3, PI_4, PI_5, PI_7$) after embedding the shuffled code. These changes, from (7,6,4,6,5) to (6,7,5,5,4), could potentially cause errors during LBP extraction. The issue arises because LBP relies on the relationship between a central pixel's intensity (P_c) and its neighbors (v). Embedding a bit "0" in PI_1 (originally 7) would decrease its value to 6 (PI_1'), violating the relationship with Pc (previously >= 7). Similarly, embedding a bit "1" in PI_3(originally 6) would increase it to 7 (PI_3'), again breaking the relationship. While adjusting the values by +-1 might seem like a solution, it would introduce errors during bit extraction (extracting 1 from

$PI_1(7)$ and 0 from $PI_3(6)$). To address this and preserve the embedded bit, we opted for a ± 2 adjustment strategy (e.g., PI_1' becomes 8 and PI_3' becomes 5), ensuring both valid LBP extraction and accurate bit retrieval.

10. Merging Blocks:
    - The modified block PI′ and the fractional block PF are merged to create the watermarked block P^*.
11. Inverse Transformation:
    - The modified LL sub-band is divided by 10, followed by applying the inverse DWT to the modified LL along with the original detail parameters (HL, LH, HH) to produce the Normalized Watermarked Medical Image (*NWMI*).
12. Final Watermarked Image:
    - Finally, the watermarked image is generated by denormalizing the *NWMI* to obtain the Watermarked Medical Image (*WMI*).

Fig. 26 illustrates the watermark embedding process, providing a visual reference for the steps outlined.

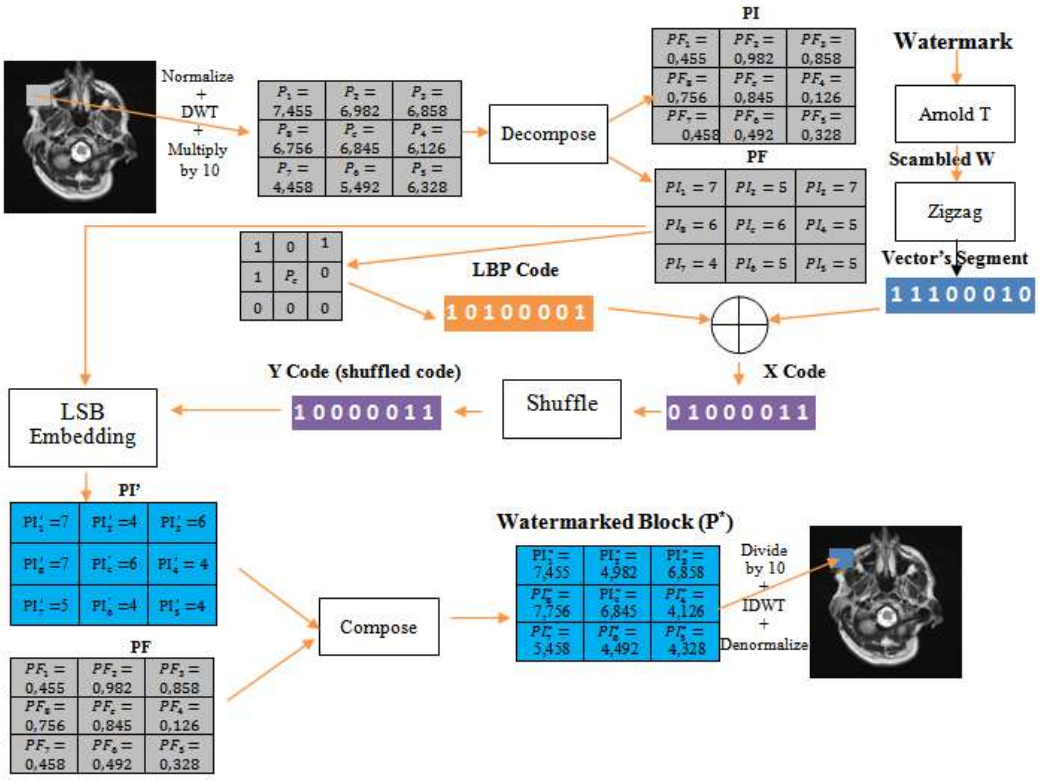


Fig 26. Embedding algorithm of the proposed method.

IV.2.4. Extraction Process

The watermark extraction process operates in reverse to the embedding process, allowing for the recovery of the original watermark from the watermarked image. The steps involved in the extraction process are as follows:

1. Normalization and Transformation:
    - The watermarked image is normalized, and a DWT is applied to retrieve the sub-band LL. The image is then multiplied by 10 to prepare it for further processing.
2. Block Division**:**
    - The watermarked sub-band is divided into non-overlapping blocks P_i of size 3x3.
3. Block Splitting**:**
    - Each block *P* is split into two parts: *PI* (containing integer values) and *PF* (holding fractional values).
4. Calculation of Shuffled Code**:**
    - For each block PI, the 8 bits of the LBP code are extracted using equations (1) and (2).
    - The Least Significant Bits (LSB) are extracted from the coefficients of the modified sub-band by analyzing the eight coefficients surrounding the central coefficient to obtain the shuffled code Y_j.
5. Recovery of Original Bits**:**
    - The pixel values X_j are recovered from the shuffled code Y_j using the following equation:

$$X_j=\{ x_j^0 = y_j^1; x_j^1 = y_j^0; x_j^2 = y_j^3; x_j^3 = y_j^2; x_j^4 = y_j^5; x_j^5 = y_j^4; x_j^6 = y_j^7; x_j^7 = y_j^8 \} \quad (14)$$

6. Extraction of Watermark Bits**:**
    - An XOR operation is performed between the X_j code and the LBP code to extract 8 physical bits from the watermark sequence:

$$W_i^A = X_j \oplus LBP_j \quad (15)$$

7. Concatenation of Watermark Segments**:**
    - The extracted bits from each watermark segment are concatenated to form the complete watermark vector.
8. Inverse Zigzag Scan**:**
    - The inverse Zigzag scan operation is applied to the modified vector, as performed in the fifth embedding step, to reconstruct the matrix of modified LL.
9. Inverse Arnold Transform**:**
    - Finally, the inverse Arnold transform is applied to the scrambled watermark W^A using the secret key to obtain the extracted watermark *EW*.

Fig. 27 provides a graphic depiction of the watermark extraction procedure, illustrating each step of the process.

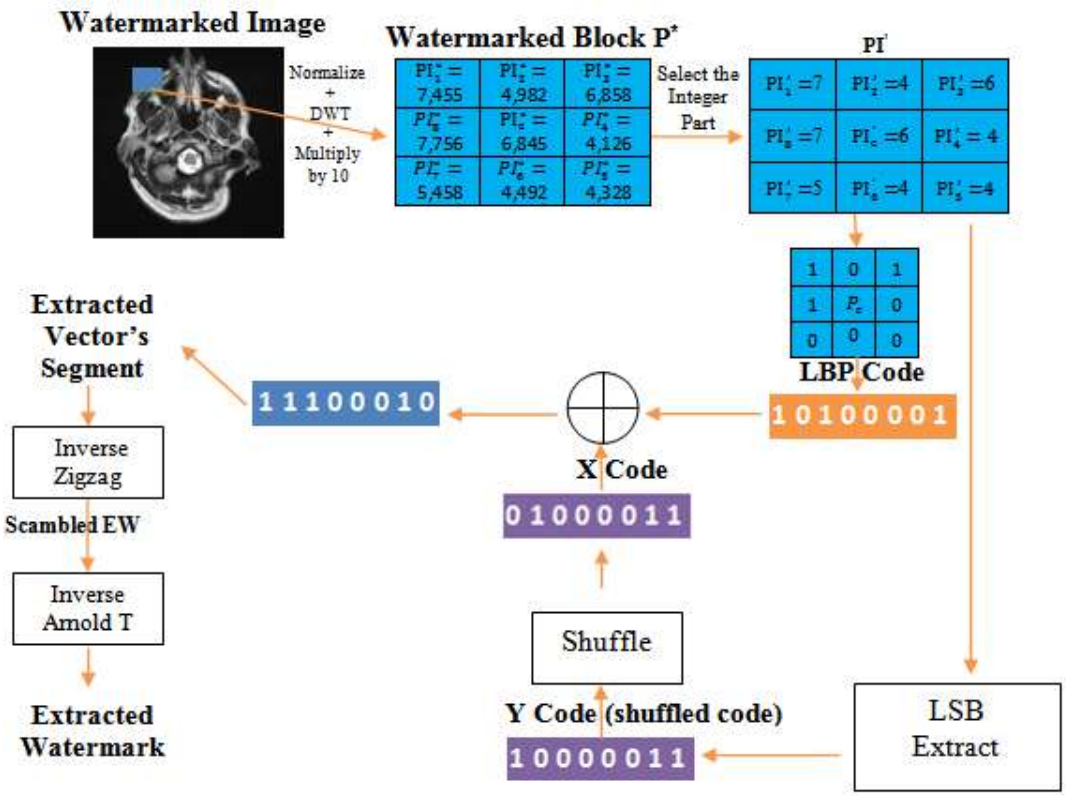


Fig 27. The process of extracting the watermark's eight bits

IV.3. Medical Image Watermarking Based on gradient analyses and DWT

In watermarking, enhancing robustness against noise is crucial to ensuring the integrity of the embedded watermark. One way to achieve this is through binning on the gradient direction. This process sorts gradient directions into specific bins, allowing us to handle variations in pixel intensity more effectively. The following outline this approach.

IV.3.1. Background Concepts

To understand this approach in detail, it is important to review the underlying background concepts**:**

a) Gradient

In image processing, the image gradient plays a crucial role. It helps us identify how the intensity (brightness) of an image changes in different directions. The gradient is characterized by two key aspects: magnitude and direction. Magnitude tells us the strength of the change, while direction indicates where the change is happening. Essentially, each pixel in the gradient image represents the rate of change in the original image's pixel value at that specific location, relative to a particular direction [54].

Mathematically, these properties can be calculated using the following equations:

$$\mathrm{Mag} = \sqrt{g_y^2 + g_x^2} \qquad (16)$$

$$\lambda = \tan^{-1}\left(\frac{g_y}{g_x}\right) \qquad (17)$$

Where the gradients in the y and x directions are denoted by gy and gx, respectively.

Mag represents the gradient's magnitude. λ is a measure of the gradient's direction.

IV.3.2. Embedding process

Step 1. In this step,we first normalize the medical image and converts it to the frequency domain by DWT to extract the sub-bands (LL,LH,HL,HH), we choose LL sub-band.

Step 2. Compute the gradient of the LL subband, including both the magnitude and direction.

Step 3. Divide the LL image, gradient magnitude and direction into 3 × 3 nonoverlapping blocks(*P*). The nonoverlapping blocks prevent data loss.

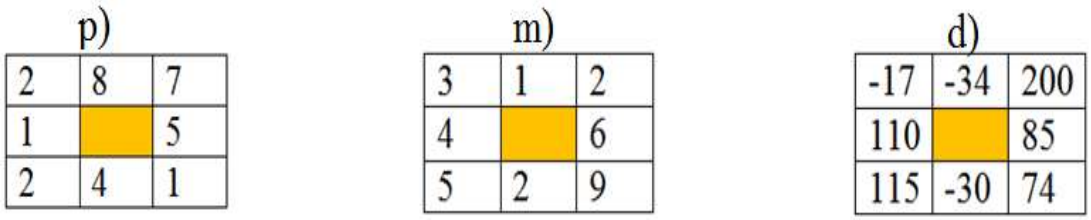


Fig 28. Illustration of blocks p) Block of original image, m)Block of gradient magnitude d)Block of gradient direction

Step 4. The continuous gradient direction values are binned into discrete intervals. This binning process ensures that small changes in direction (due to noise or image attacks) do not significantly affect the watermark. The binning formula is:

$$f(x)=\lfloor \frac{x}{c} \rfloor \text{ for } 0 < x \leq 2\pi \qquad (18)$$

Where ⌊.⌋The ceiling function, which returns the smallest integer greater than or equal to the input value.

In this work, the constant $c=\frac{\pi}{2}$ divides the gradient direction into four intervals, allowing flexibility in aligning the watermark with the image's directional structure, as shown in Fig. 29.

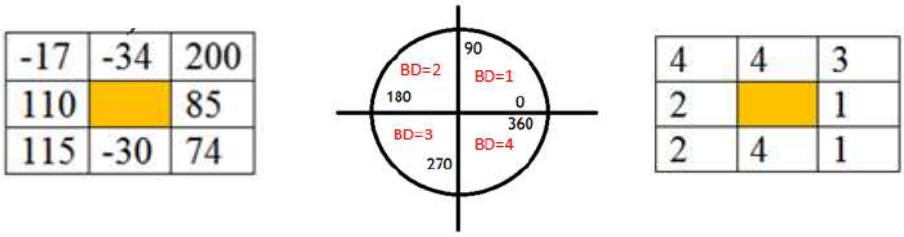


Fig 29. The block transformation process: a) the original gradient direction block, b) the decomposition circle, and c) the transformed gradient direction

Step 5. The gradient magnitudes are binarized based on a threshold value α, which is set as the midpoint of the range of gradient magnitudes in each block. The binarization function is defined as:

$$B\alpha(M(i,j))=\begin{cases}1 & \text{if} \quad M(i,j) \geq \alpha \\ 0 & \text{if} \quad M(i,j) < \alpha\end{cases} \qquad (19)$$

This step creates a binary representation (as shown in Fig. 30) of the gradient magnitudes, which will later be used to generate the crypto_code.

0	0	0
1		1
1	0	1

Fig 30. Representation of B_α

Step 6. Construct the watermark final histogram H accumulating binarized magnitudes according to the binned directions as follows:

$$H[k] = \sum_{d \in D} \sum_{(i,j) \in Skd} B\alpha(M(i,j)) \qquad (20)$$

Where D is the set of all possible binarized directions, S_{kd} is the set of all pixel indices (i, j) having gradient directions d and fall into the k^{th} bin of the histogram.

Step 7. For each block, a binary code called the crypto_code is generated by accumulating the binarized gradient magnitudes corresponding to the binned gradient directions. The crypto_code consists of 8 bits, which is the exact length needed to encrypt each segment of the watermark.

If the length of the crypto_code is more or less than 8 bits, it is adjusted uniformly across all blocks by either trimming or padding with additional bits.

Step 8. The watermark, divided into segments of 8 bits, is encrypted using the crypto_code generated for each block. The XOR operation is applied to combine the watermark bits and the crypto_code:

$$X_j = W_i^A \oplus \text{crypto\_code}_j \qquad (21)$$

This step ensures that the watermark is securely embedded and can only be extracted using the correct crypto_code.

• **Step 9.** The Least Significant Bits (LSB) of the coefficients of the watermarked *LL* sub-band are replaced by the encrypted watermark bits. This subtle modification ensures that the watermark is embedded in a way that is imperceptible to human vision.

Step 10. After modifying the LL sub-band with the embedded watermark, an inverse DWT is applied. The modified LL sub-band is combined with the unchanged LH, HL, and HH sub-bands to reconstruct the watermarked image.

Fig. 31 illustrates the embedding process of this method.

IV.3.3. Extraction process

We carry out the same embedding steps from step 1 to step 6, then follow the following steps:

Step 1. The watermarked image is decomposed into sub-bands (LL, LH, HL, HH) using DWT.

Step 2. The modified LL sub-band is divided into non-overlapping blocks of size n ×n.

Step 3. The Least Significant Bits (LSBs) are extracted from each block's coefficients to retrieve the encrypted watermark segments.

Step 4. Using the same crypto_code (constructed from the gradient information of the LL sub-band), the encrypted watermark segments are decrypted with an XOR operation.

$$W_i^A = Y_j \oplus \text{crypto\_code}_j \tag{22}$$

Step 5. The decrypted segments are concatenated to form the original watermark.

Fig. 32 illustrates the extraction process of this method.

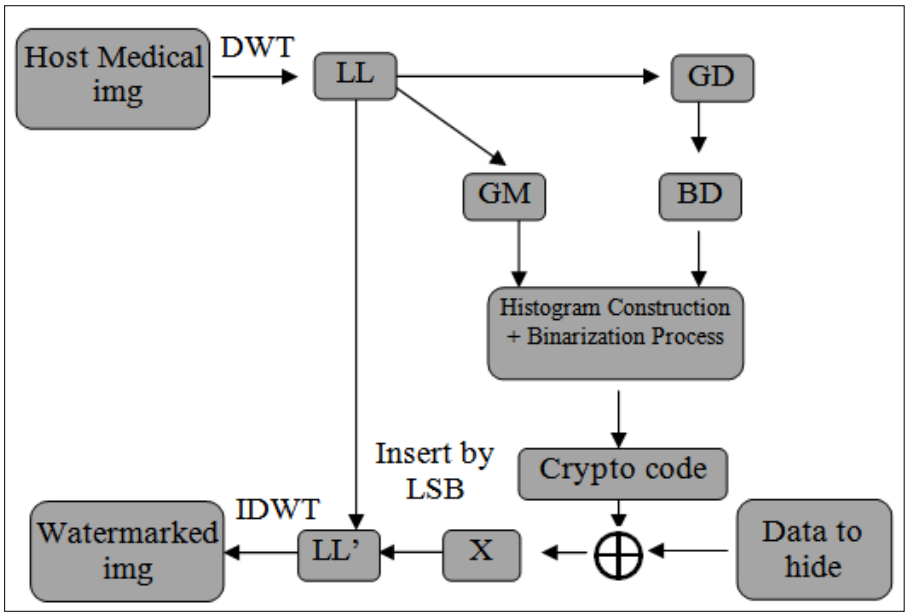


Fig 31.Watermark embedding, GM: Block (3x3) of gradient magnitude, GD: Block (3x3) of gradient direction, BD: Binning Decomposition

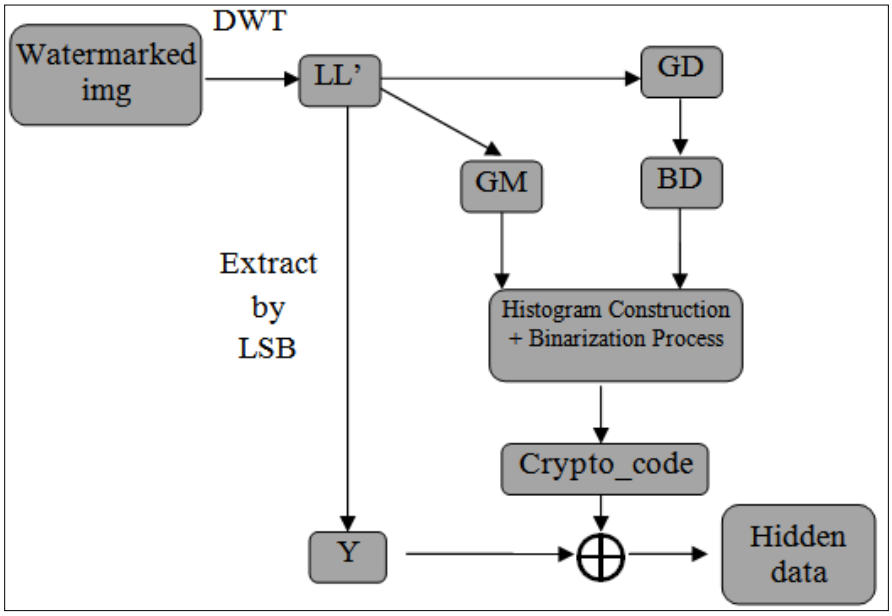


Fig 32. Watermark extraction, GM: Block (3x3) of gradient magnitude, GD: Block (3x3) of gradient direction, BD: Binning Decomposition

IV.4. Zero Medical Image Watermarking Based VGG16

This sub-section introduces a zero-watermarking method designed specifically for medical images. Unlike traditional zero-watermarking techniques that rely on manual feature extraction, the proposed method combines deep learning with zero-watermarking to achieve superior robustness. The core of this method lies in the use of a pre-trained VGG16 model for feature extraction, along with secure watermark encryption and decryption techniques.

IV.4.1. Overview of the Method

The key components of the proposed method include:

1. VGG16 for Feature Extraction
2. Watermark Encryption and Decryption
3. Encrypted Watermark Generation and Extraction

These components work together to create a robust watermarking process that can withstand various attacks while ensuring the watermark remains imperceptible in the medical image.

IV.4.2. VGG16 for Feature Extraction

Step 1: Pre-trained VGG16 Model

- Goal: Leverage the feature extraction capabilities of a deep learning model for medical images.
- Procedure:

  The VGG16 network, pre-trained on ImageNet, is used as the backbone of the proposed algorithm. VGG16 is a Convolutional Neural Network (CNN) known for its excellent feature extraction abilities. In this method, VGG16 extracts high-level semantic features from the medical image, which are then used in the zero-watermarking process.

Why VGG16?

The choice of VGG16 is based on its depth and architecture, which can capture complex patterns and details in images. For medical images, these features are crucial for

embedding a robust watermark that is resistant to distortions such as noise, compression, and attacks.

Fig. 33 illustrates the architecture of the VGG16 model.

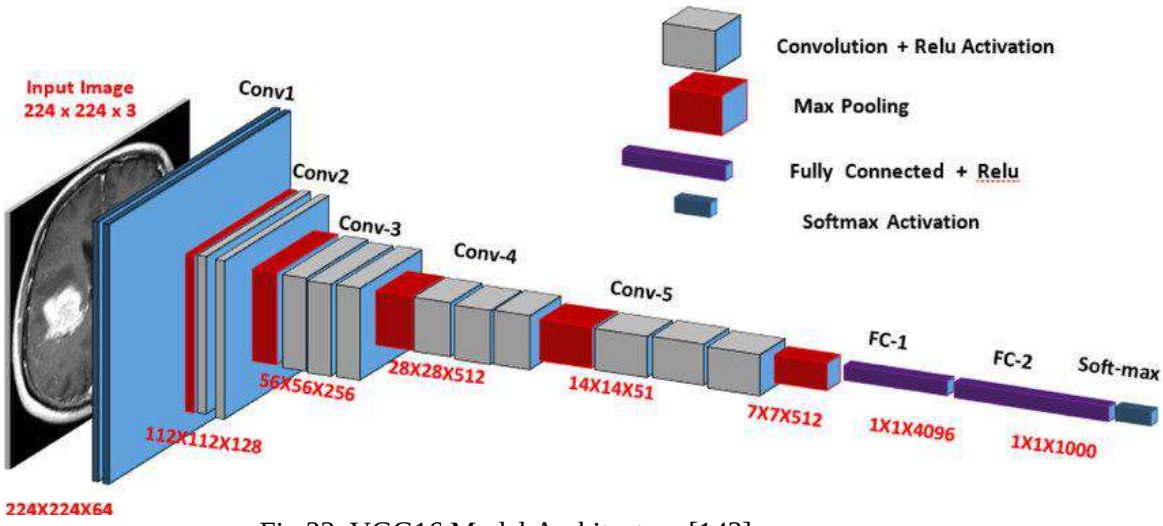


Fig 33. VGG16 Model Architecture [143].

Step 2: Feature Extraction Process

- Input: The medical image (MI) is resized and fed into the pre-trained VGG16 model.
- Output: The deep features extracted from the convolutional layers. These features represent significant image content, such as texture, edges, and shapes.

The extracted features are obtained from a specific layer of VGG16 (commonly one of the middle convolutional layers, In our experiment we chose block5_conv3 ,the layers are shown in Fig. 34**)** since these layers capture meaningful patterns without being too abstract or too basic.

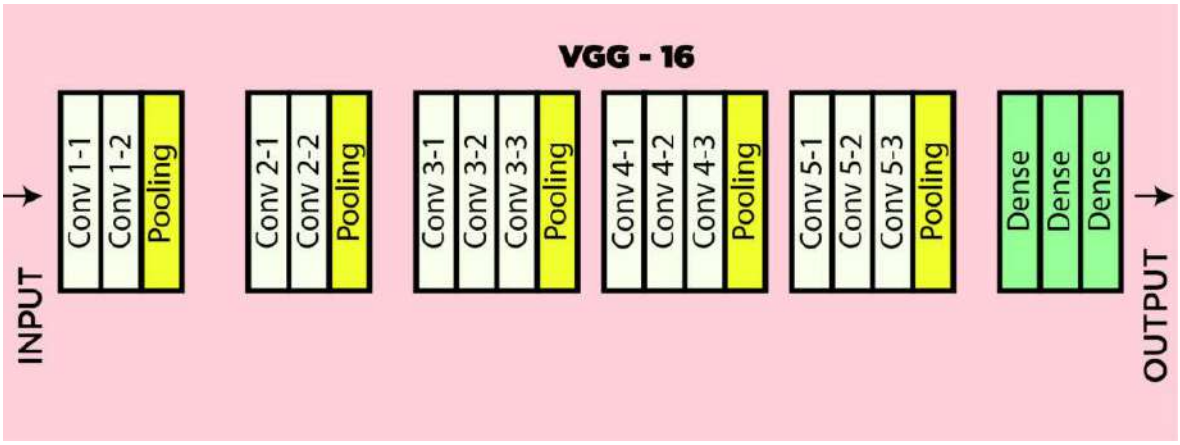


Fig 34. VGG-16 architecture Map[143]

Step 3: Feature Selection

- Goal: Select the most relevant features for watermark embedding.
- Procedure:

  Once the features are extracted from VGG16, the features corresponding to the most significant image regions (such as edges or textures) are selected for watermark generation. This ensures that the watermark is embedded in regions where attacks (e.g., compression or noise) are least likely to affect the image.

IV.4.3. Watermark Encryption and Decryption

Step 1: Watermark Representation

- Goal: Prepare the watermark for secure embedding.
- Procedure:

  The watermark, typically a binary image 64x64 (e.g., logo), is transformed into a sequence of bits. To ensure security and robustness, this binary sequence is encrypted before being combined with the medical image features.

Fig 35. Exemple of Watermark

Step 2: Encryption Process

- Goal: Encrypt the watermark to prevent unauthorized access or tampering.
- Procedure:

  The encryption process uses a secure cryptographic technique such as Arnold Transform. The binary sequence of the watermark is encrypted using a secret key, producing an encrypted watermark.

  The encryption function can be represented as:

$$We=Encrypt(W,K) \qquad (23)$$

  where *We* is the encrypted watermark, *W* is the original binary watermark, and K is the secret key.

  An example of watermark encryption using the Arnold Transform is illustrated in Fig. 36, with parameters a=1, b=2, and a key (number of iterations) set to 10, as specified in Eq. 10.

Fig 36.Illustration of watermark and encryption watermark

IV.4.4. Watermark Embedding Process

Step 1: Feature-Watermark Mapping

- Goal: Generate the Master Share based on the deep features and encrypted watermark.
- Procedure:

  The encrypted watermark is not directly embedded into the medical image. Instead, the deep features extracted from VGG16 are combined with the encrypted

watermark to generate a Master Share. This process ensures that the watermark does not alter the image itself but is linked to its deep features.

The feature vectors extracted from VGG16 are segmented to match the size of the encrypted watermark bits. The encrypted bits are then mapped to the corresponding segments of the feature vectors.

Step 2 Master Share Generation

- Goal: Create a master share that ties the watermark to the image's deep features.
- Procedure:

  The final a master share is generated by xoring the encrypted watermark bits with the most significant feature vectors extracted from the VGG16 model. This master share is a unique identifier that can be used to verify the integrity and ownership of the medical image.

IV.4.5. Watermark Extraction Process

Step 1: Feature Extraction

- Goal: Re-extract the deep features from the potentially altered medical image.
- Procedure:

  During watermark extraction, the VGG16 model is once again used to extract the deep features from the medical image. These features are compared with the original deep features used during the embedding phase.

Step 2: Encrypted Watermark Recovery

- Goal: Recover the encrypted watermark by reversing the watermark mapping process.
- Procedure:

The encrypted watermark bits is generated by xoring the master share with the most significant feature from the medical image. Since the watermark is linked to the image's features, even if the image undergoes minor distortions, the encrypted watermark can still be recovered.

Step 3: Decryption of Watermark

- Goal: Retrieve the original watermark.
- Procedure:

  The encrypted watermark is decrypted using the same secret key that was used during the embedding phase. The decryption function is:

$$W=\text{Decrypt}(We,K) \quad (24)$$

  The recovered watermark is then compared with the original watermark to verify its authenticity.

IV.4.6. Proprieties and Advantages

The proposed method offers several key advantages:

- Deep Feature-Based Embedding: Using VGG16 for feature extraction makes the method highly robust to various attacks, as deep features are less susceptible to changes in pixel-level details.
- Zero-Watermarking: Since the watermark is not directly embedded in the image, the integrity of the medical image is preserved. This is crucial for applications in telemedicine and medical image storage, where any alteration to the image can affect diagnoses.
- Secure Watermarking: The encryption of the watermark ensures that it cannot be easily accessed or modified by unauthorized parties.
- Robustness: The method is designed to resist a wide range of attacks, including noise, compression, and geometric distortions, due to the combination of deep learning and zero-watermarking techniques.

Fig. 37 presents a detailed depiction of the embedding and extraction process.

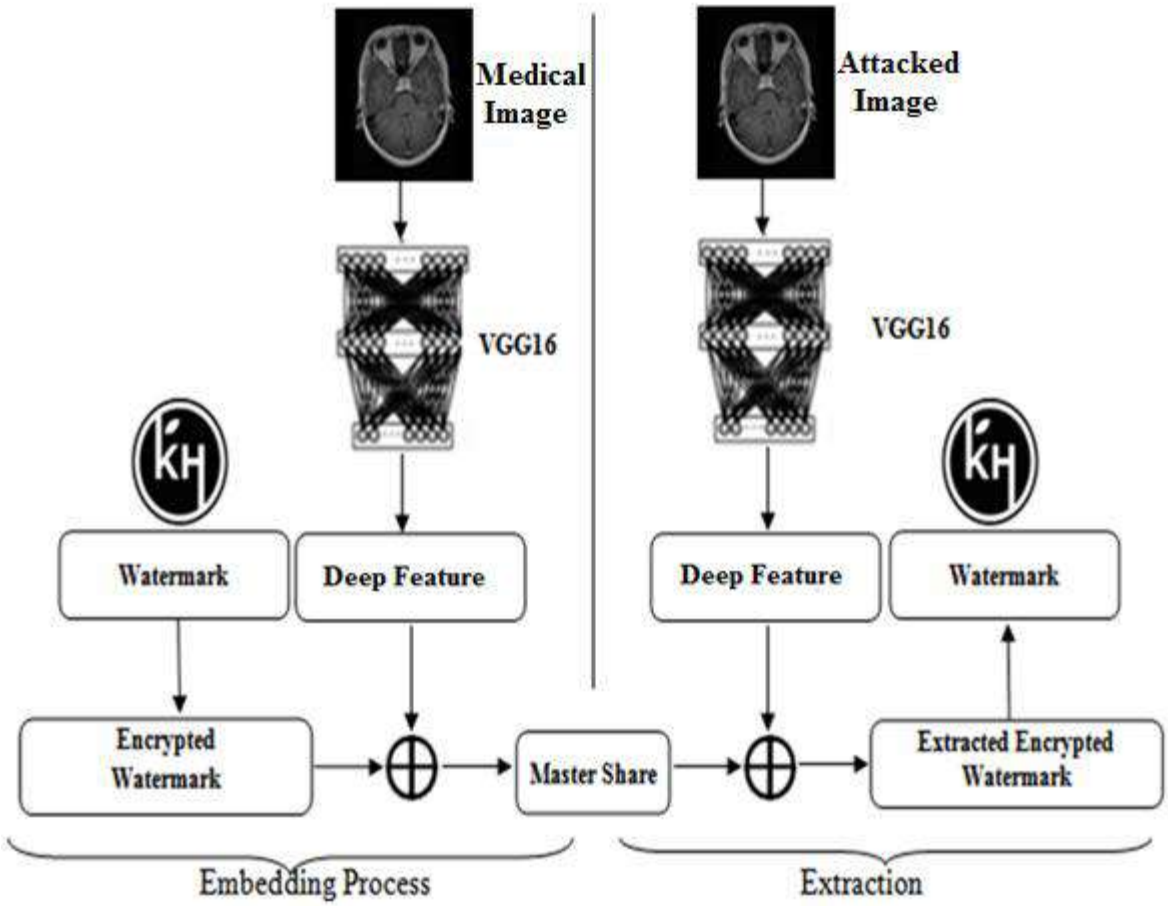


Fig 37.The process of the proposed method

IV.5. Conclusion

This chapter introduced three key methods for medical image watermarking. The first, LBP-DWT-based blind watermarking, combined Local Binary Patterns and Discrete Wavelet Transform to ensure robustness against various attacks. The second, DWT with Gradient Analysis, leveraged gradient magnitudes and directions for enhanced security, especially against geometric distortions. The third, a Deep Learning-Based Zero-Watermarking method using a pre-trained VGG16 model, provided strong adaptability and robustness against complex attacks. Each method contributes to advancing secure and reliable medical image watermarking, addressing different challenges in image protection and data authentication.

V. Chapter 4 : Experimental Results and Discussion

V.1. Introduction

This chapter evaluates the proposed medical image watermarking methods, focusing on their robustness, imperceptibility, and security under various conditions. It assesses how well watermarks can be embedded and extracted in the presence of image distortions and attacks, while preserving the diagnostic quality of medical images. Methods are tested using performance metrics like PSNR for image quality, NC and BER for extraction accuracy, and SSIM for perceptual integrity. A range of attacks, including noise, blurring, and compression, are applied to validate the methods' resistance. The chapter highlights the strengths and trade-offs between handcrafted and deep learning approaches.

V.2. First approach: watermarking based on DWT and LBP

V.2.1. Experimental setup

To evaluate the proposed watermarking scheme, a diverse set of medical image types has been utilized, including X-rays, CT scans, ultrasounds (US), and MRI images. This variety ensures that the effectiveness of the watermarking methods can be assessed across different imaging modalities, each with its unique characteristics and challenges. For clarity in analysis, alphanumeric labels have been assigned to represent the image names, facilitating easy reference throughout the experiments. All medical images are standardized to a resolution of 512×512 pixels, as illustrated in Fig**.** 38, which shows a total of twelve (12) samples.

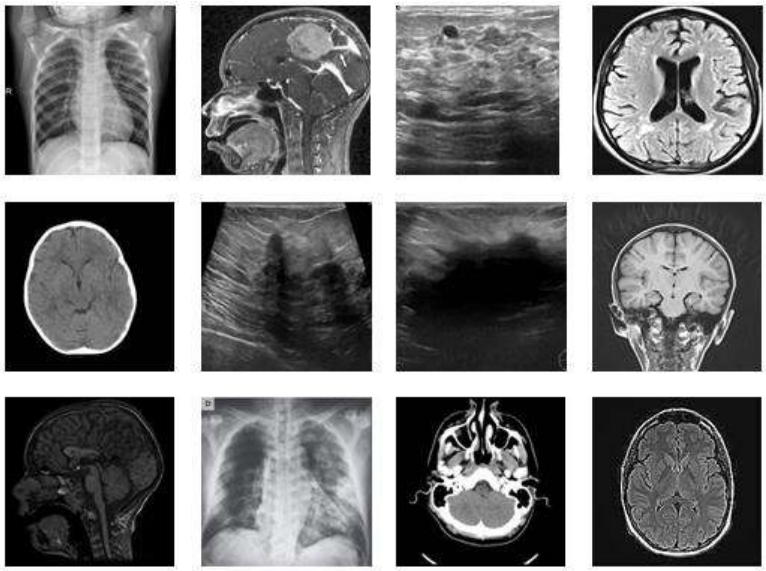


Fig 38. Fig 17 .X-ray, CT, US, and MRI host medical images utilized in our suggested method

In our experiments, the watermark is derived from the patient's information, which has been converted into a binary image format. This text-based watermark is commonly employed for copyright and ownership purposes, ensuring that sensitive information is securely embedded within the medical images without compromising their integrity (see Fig**.** 39). The use of binary images for watermarking allows for effective embedding while maintaining the imperceptibility of the watermark, which is crucial in medical applications.

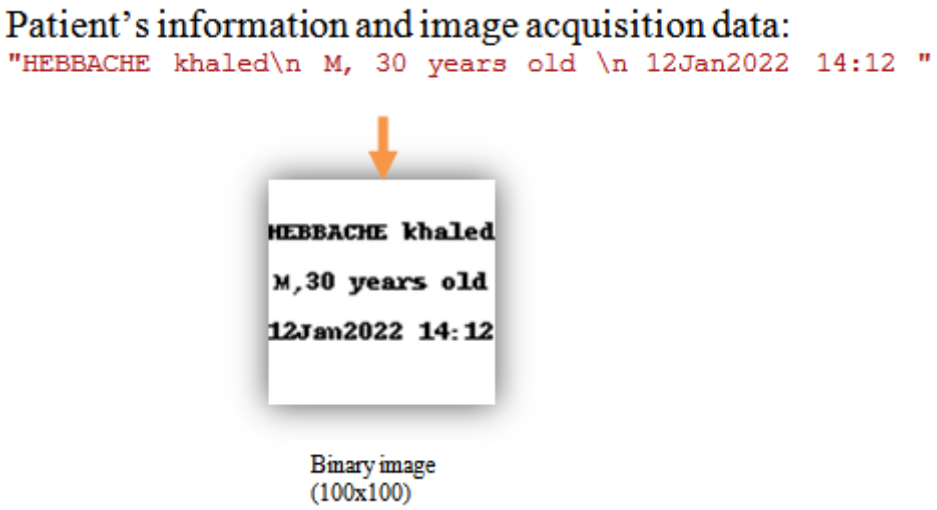


Fig 39.Creation of watermark

The images utilized in this study were taken from various Kaggle datasets, providing a rich repository of high-quality medical images, ensuring the relevance and applicability of the experiments in real-world scenarios. The experimentation was conducted using Google Colab software, enabling efficient execution of the watermarking algorithms and the application of various image processing techniques.

V.2.2. Experimental setup

To ensure robustness, we consider using the DWT in the frequency domain along with the LBP in the spatial domain to maintain the imperceptibility. This approach effectively blends the spatial and frequency domains to balance the trade-off between robustness and imperceptibility. Fig 40 presents the outcomes of using the system against CT, MRI, X-ray, and US scans.

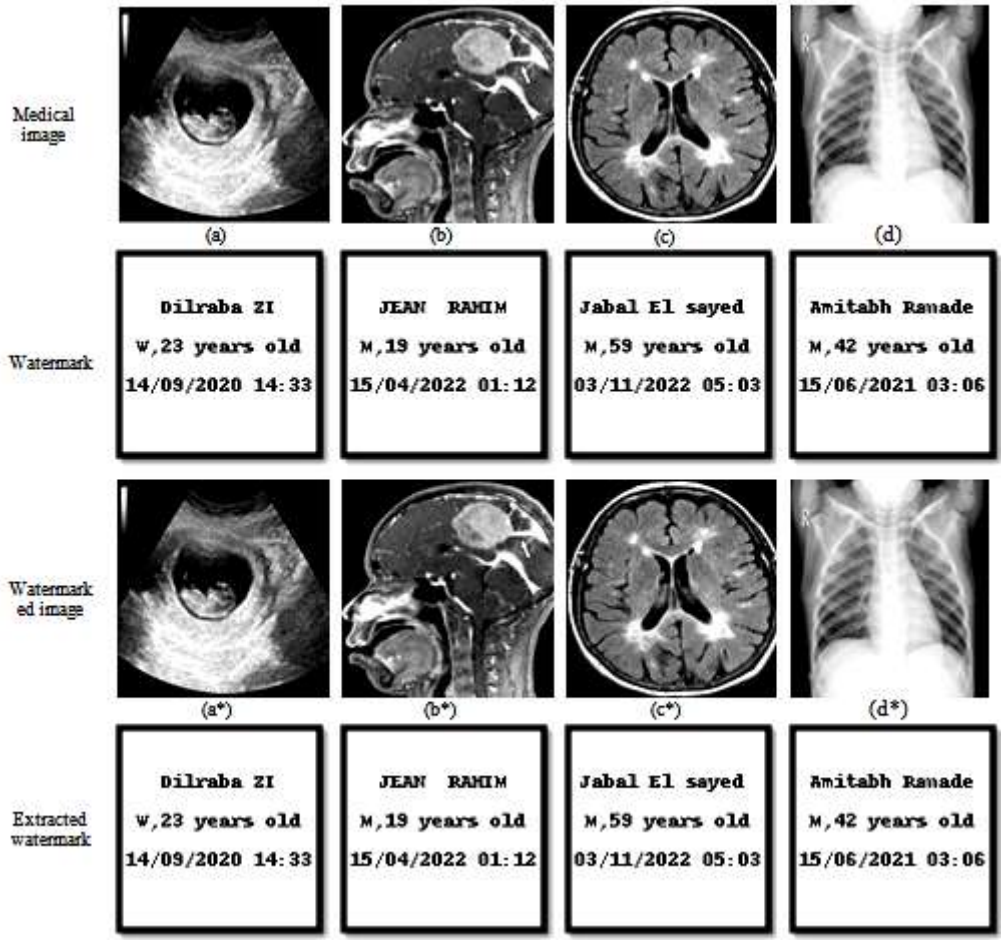


Fig 40.Exemple illustrate watermarked MI and its extracted watermark

a) **Imperceptibility Evaluation**

Medical image watermarking must carefully maintain image quality while integrating sensitive patient data. This is because watermark insertion process may introduce distortions to the medical image. Thus, to assess the imperceptibility of our suggested approach, we measure the image quality after inserting the watermark using the metrics defined in section 5.1 i.e., PSNR and SSIM.

Fig 41 presents the impartibility results in terms of PSNR and SSIM.

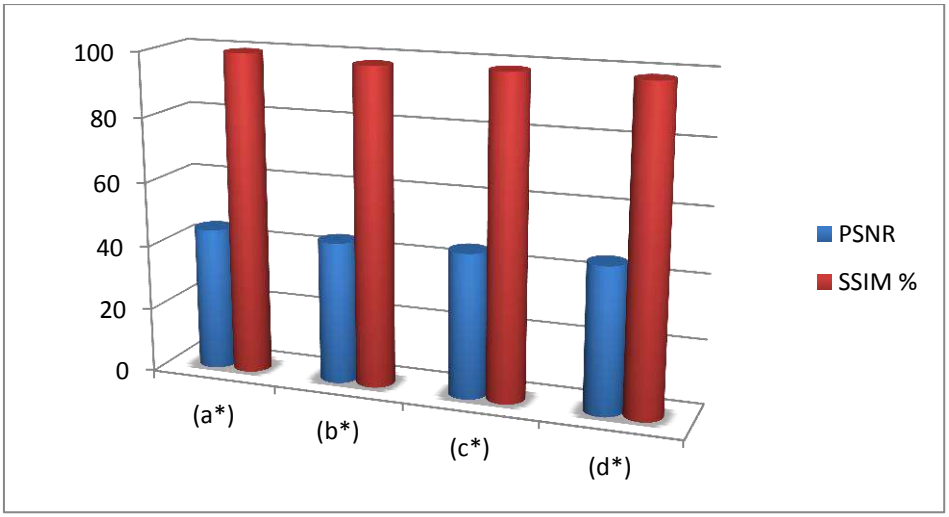


Fig 41. Imperceptibility results in terms of PSNR and SSIM.

From the results we obtained, it is clear that the PSNR and SSIM values significantly exceed the threshold of 44 dB and 98%, respectively. This means that embedding the LBP code using LSB does not affect the image and keeps the imperceptibility value high. These high values indicate that the original *MIs* and their corresponding *WMI*s show close visual similarity .This indicates that the visual systems of humans are unable to distinguish differences between host *MI* and their corresponding *WMI* due to their extreme similarity. These high PSNR and SSIM values affirm that the watermarking techniques preserve the visual quality of *MIs* with an imperceptible impact, which is crucial for maintaining the integrity of diagnostic images in the medical field.

b) **Robustness tests**

In the absence of attack, we use the suggested extraction algorithm to extract watermarks from watermarked medical images. Fig 40 show these extracted watermark from watermarked images including CT scan, MRI, X-Ray and Ultra-sound, and all results are with NC=1 and BCR = 1. This denotes the precise and accurate extraction of the watermark. So, we can extract accurately watermark images in no attack case. In order to preserve the patient's information, the watermarking algorithm must resist to the various attacks. In order to verify the robustness of the proposed algorithms, we test our algorithm against various types of attacks, including JPEG compression, Gaussian noise, Salt and Pepper noise, median filter and other attacks. The NC and BCR are then calculated between the original and

extracted watermark. In the next subsections, we provide more details on the performance of our method against different kinds of attacks.

- **Robustness against image compression**

In Fig 42 illustrates the BCR and NC values for the watermarks extracted after subjecting them to a JPEG compression attack. In this case, the quality factor (Q) was adjusted within the range of 10 to 70[47]. From Fig 42, it appears that, for all the types of medical images, the watermark is successfully restored under strong JPEG attacks (with BCR =1 and NC = 1) when the quality factor (Q) exceeds 70. Fig 43 shows how the extracted watermark transforms as the factor (Q) changes in response to a JPEG compression attack.

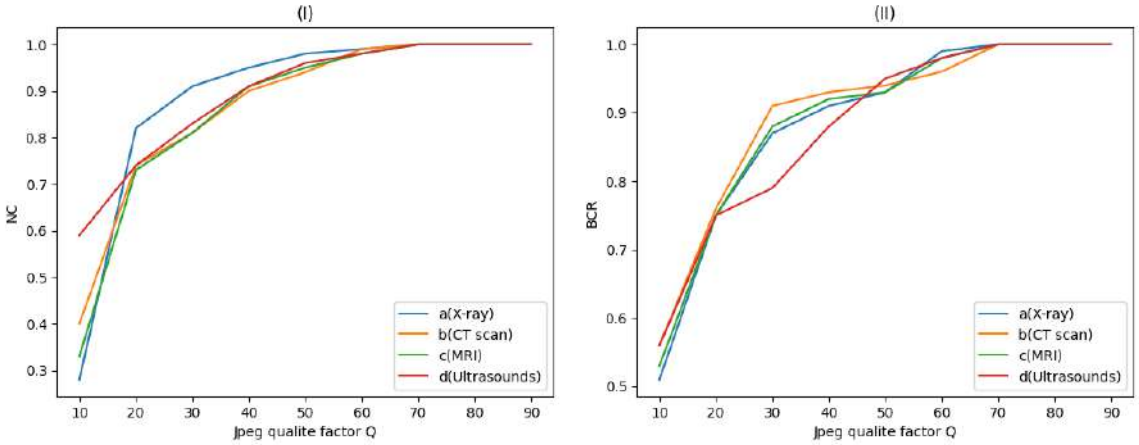


Fig 42 .Robustness against compression attacks. (I) NC curve by Q . (II) BCR curve by Q

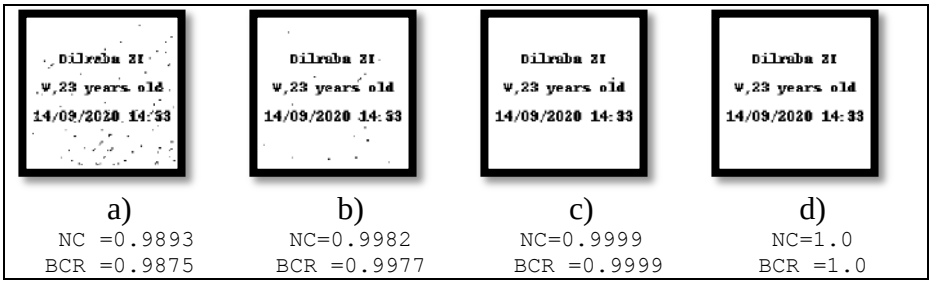


Fig 43. The extracted watermark according Q: a) Q=50,b)Q=60,c)Q=70,d)Q=80

Despite the change in the value of Q, we can note that the readability of the extracted watermark initiates at Q=50, as shown by Fig 43. Furthermore, Fig 42 indicates that both NC and BCR tend to approach a value of 1 when Q surpasses 60. These observations from the

two figures affirm the robustness of our method when subjecting a watermarked medical image to a JPEG compression attack.

- **Robustness against image processing attacks**

In the following analysis, we assess the effectiveness of the proposed method when subjected to various image processing attacks, including noise addition, low pass filtering (Average, Gaussian, Median), gamma correction, histogram equalization, Laplacian sharpening, and bit-plane removal. In the bit-plane removal attack, the least significant bits of the watermarked image are substituted with zeros.

Table displays the BCR and NC values for the extracted watermarks after these attacks. While smoothing filters like average and Gaussian filters (3x3) have minimal impact (BCR and NC values remain high across all image types), a larger Gaussian filter (5x5) reduces effectiveness for CT scans (BCR drops from 0.9760 to 0.9367) and ultrasounds. Similar trends hold for the median filter. Salt and pepper noise has a moderate effect – low levels (variance = 0.01) cause minimal change, while higher levels (variance = 0.02) lead to a decrease in BCR and NC (e.g., X-ray NC falls from 0.9679 to 0.9141). Histogram equalization poses a significant challenge, especially for CT scans and X-rays (BCR drops to around 0.2) due to its impact on intensity distribution. Sharpening filters like Laplacian sharpening also negatively affect CT scans and ultrasounds (BCR drops below 0.7), likely due to modifications in image edges. Interestingly, MRIs display surprising resilience in some cases. Removing a moderate number of bit planes (5 bits) has a moderate impact on CT scans (BCR falls to 0.9575) but minimal effect on MRIs (BCR remains around 0.9). Overall, the interplay between attack type, image modality, and watermark robustness is evident.

Table 7. BCR and NC values of the proposed technique under image processing attacks Fig 44 demonstrates the successful extraction of the watermark from the watermarked medical images subjected to the different attacks namely Gamma correction (gamma=1), Salt and pepper noise (var = 0.01), and Bit plane removal (plan =6). The readability of the extracted watermark highlights the robustness of our algorithms against these attacks.

Table 7. BCR and NC values of the proposed technique under image processing attacks

Attacks	A(CT scan)		B(X-ray)		C(MRI)		D (Ultrasounds)	
	NC	BCR	NC	BCR	NC	BCR	NC	BCR
Average filter (3 x3)	0.9678	0.9533	0.8177	0.7866	0.9687	0.9545	0.9353	0.9169
Gaussian filter (3 x 3)	0.9833	0.9760	0.9003	0.8792	0.9800	0.9817	0.9646	0.9580
Gaussian filter (5 x 5)	0.9524	0.9367	0.7761	0.7397	0.9575	0.9421	0.9199	0.8987
Median filter (3 x 3)	0.9666	0.9523	0.9108	0.8794	0.9174	0.9265	0.9028	0.9094
Salt and pepper noise(var = 0.01)	0.9691	0.9636	0.9679	0.9646	0.9524	0.9524	0.9703	0.9696
Salt and pepper noise(var = 0.02)	0.9278	0.9252	0.9141	0.9202	0.8916	0.8862	0.8290	0.8276
Histogram equalization	0.7195	0.2143	0.3528	0.3093	0.9433	0.5920	0.8378	0.5474
Laplacian Sharpening	0.7264	0.5854	0.5217	0.5000	0.9366	0.7551	0.8720	0.7251
Bit-plane removal (5 bits)	0. 9575	0.9551	0.4195	0.4250	0.9178	0.9238	0.8349	0.8381
Bit-plane removal (6 bits)	0.9737	0.9609	0.6710	0.6392	0.9737	0.9536	0.9362	0.9019
Gamma correction (1)	1.0	1.0	1.0	1.0	0.9833	0.9902	0.9900	0.9941

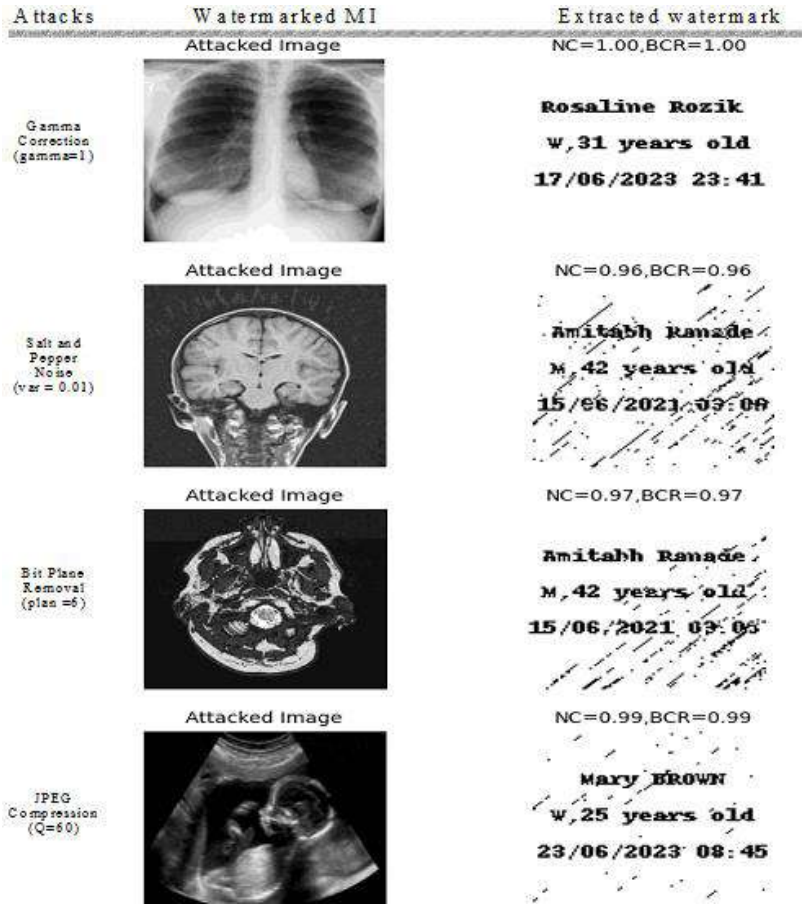


Fig 44.Attacked images with salt and pepper noise (var = 0.01), LSB removal (6 bits) and JPEG lossy (Q = 60).

- **Robustness against geometrical attacks**

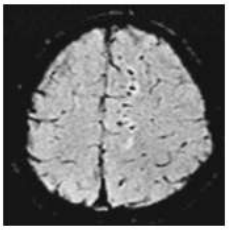
The resilience against the geometrical attacks is demonstrated in the following experiments. The suggested approach is comparatively some resistant to geometrical attacks, as shown by **Erreur ! Référence non valide pour un signet.** and Fig 45. All modalities demonstrate good robustness to resizing (512 → 256 → 512), with NC and BCR values generally above 0.88. Robustness significantly decreases with increasing rotation small angles. Performance declines with increasing rotation angles, while small rotations (0.5°) resulted in a moderate performance drop, larger rotations (1°) led to a substantial decline in both BCR and NC for all modalities. CT scans consistently demonstrated the highest resilience to rotation, while Ultrasound images proved most susceptible. This highlights the vulnerability of the current watermarking scheme to larger rotational distortions.

Cropping effects also vary – CT scans and ultrasounds generally show lower BCR and NC (around 0.7) with surrounding crops, while X-rays (> 0.9) and MRIs (around 1.0) in some cases (top-left cropping) exhibit minimal impact. These findings emphasize the interplay between attack type, image modality, and watermark robustness.

Table 8. The suggested technique's BCR and NC values under geometrical attacks

Attack	A(CT scan)		B(X-ray)		C(MRI)		D (Ultrasounds)	
	NC	BCR	NC	BCR	NC	BCR	NC	BCR
Resizing (512 → 256 → 512)	0.9787	0.9744	0.9108	0.8884	0.9128	0.9114	0.8949	0.8896
Rotation (0.5°)	0.7804	0.6887	0.6686	0.6037	0.6314	0.6519	0.6343	0.6584
Rotation (1°)	0.7045	0.6143	0.5678	0.5617	0.5734	0.5314	0.5509	0.6097
Surrounding crop (10%)	0.7462	0.6924	0.9754	0.6968	0.9950	0.7195	0.9917	0.7166
Top left quarter crop (0.25)	0.9487	0.7827	1.0000	0.9846	1.0000	0.7864	0.9996	0.7864

Top left crop (0.20)



Rotation (0.5°)

Fig 45. Illustration of some types of geometrical attacks.

c) **Comparison with other methods**

In this subsection, we compare the performance of our proposed DWT-LBP watermarking method with three other existing approaches for medical image watermarking which are presented in [142], [56], [31].

Table 9.Comparisons of imperceptibility (PSNR & SSIM) among different methods

	[56]	[142]	[31]	**The proposed**
PSNR	43.28	36.00	36,54	44,23
SSIM	0.9864	0.9884	0.9961	0.9888

Table 10.Comparisons of robustness (BCR & NC) among different methods

Attacks	[56]	[142]		[31]	**The proposed**	
		BER	NC	NC	NC	BER
Salt & Pepper Noise	0.9755	0.0926	0.9969(0.001)	0.7751(0.001)	0.9649(0.01)	0,0375
Gaussian noise	0.9245	0.2015	0.9941(0.001)	0.3226 (0.05)	0.9570(0.01)	0,0513
Speckle Noise	/	/	/	0.9687 (0.0001)	0.5142 (0.0001)	0,2858
JPEG compression	0.7662 (q =20)	0.3162 (q=20)	0.9997 (q=90)	0.9904 (q=60)	0.9970 (q=60)	0,0081
Median filtering (3 x 3)	0.8735	0.2515	0.9987(2x2)	0.9457	0.9244(3x3)	0,0831
Cropping	0.8913	0.1723	/	0.9337	0.948707	0,0447
Scaling	0.7906	0.2205	0.9999(2 0.5)	0.7075(2)	0.8916(2 0.5)	0,1298
Sharpening	0.8258	0.2234	/	0.6763(0.1)	0.7641	0,3586
Histogram equalization	0.8815	0.2212	/	0.7223	0.7133	0,5803

The approach presented in [56] involves the extraction of the frequency content of the image using a redundant DWT. The final coefficients are then subjected to Schur decomposition, and the watermark bits are inserted by altering the Eigen values' LSB. In [142], the lower resolution approximation module of DWT and the approximation coefficients from LWT are selected for the watermark embedding. Based on LBP features, the scaling and embedding factors are adaptively computed. Concerning the approach presented in [31], text and image watermarks are prepared, the cover image undergoes DWT decomposition, and *HL*

and *LH* components are used to embed the medical watermark. The encoded text watermark is placed in the *HH* sub-band of the DWT coefficients.

Table 9 presents a comparison between the proposed scheme and other compared methods in terms of imperceptibility, it can be notice that the our scheme have the highest PSNR and the second highest SSIM (44,23 and 0.9888 respectively) ,we can therefore conclude from the obtained comparison our approach is guaranteeing a reasonable PSNR and SSIM values.

Table 10 reveals that the proposed method has surpassed the compared methods in terms of robustness when facing comparable attacks (the parameters for the attacks are 0.01, 0.01, 60, and 3x3 for Salt & Pepper noise, Gaussian noise, JPEG compression, and median filter, respectively, instead of 0.001, 0.001, 90, and 2x2). However, our method shows some weakness when applying the sharpening and histogram equalization attacks, which is a common weakness shared with other methods. These findings suggest that the amalgamation of the LBP and DWT domains yields better results than employing LBP exclusively within the spatial domain for medical image watermarking.

In our study, we investigate a model that integrates LBP , DWT Transform to achieve an efficient blind and robust watermarking approach for medical digital images. Our findings reveal that our method exhibits resilience to multiple attacks, primarily due to the implementation of DWT in the LL sub-band. This sub-band encapsulates diverse features from the original medical image, preserving a substantial amount of information.DWT transform generates a series of sub-bands. As a result, attacks are dispersed throughout all of the sub-bands that are generated from the DWT, and there is very little chance that the attack will alter the value of the pixel it has embedded, and then we apply LBP to the first-level LL sub-band within the wavelet transform and use it s a means to enhance the security of the watermark. While our proposed method outperforms modern approaches in several aspects, it does exhibit vulnerability to certain attacks, such as sharpening and Histogram equalization. Addressing these vulnerabilities constitutes a significant area for future research. As part of future work, we propose to explore DWT with second and third-level wavelet transforms, as well as consider the Shearlet Transform, Curvelet Transform or Contourlet Transform, all of which offer sub-band decomposition for enhanced image analysis within the frequency domain. Additionally, we can investigate the application of LBP to other sub-bands, including HL, LH, and HH, to enhance robustness against specific attack types. Our approach also

leaves room for potential enhancements, such as combining DWT with DCT or exploring other advanced transforms.

V.3. Second approach: watermarking based on gradiant analyses and DWT

V.3.1. Experimental setup:

In this section, we present a comprehensive evaluation of our proposed watermarking scheme, focusing on its robustness, imperceptibility, and computational efficiency. The evaluation encompasses a comparative analysis against related works to highlight the advantages and improvements offered by our approach. The results demonstrate the effectiveness and reliability of our watermarking method in practical scenarios.

The images utilized in our experiments, as illustrated in Fig. 46, were sourced from the COVID-19 dataset available on Kaggle (https://www.kaggle.com/datasets/fusicfenta/chest-xray-for-covid19-detection, accessed on 15 January 2024). This dataset comprises two categories of chest X-ray images: one depicting COVID-19 infections and the other representing normal lung conditions.

All images are grayscale, containing only the intensity channel, and are formatted as JPG files. The sizes of the images vary, ranging from 1200 px to 2500 px. This variability allows for a robust evaluation of the watermarking scheme across different resolutions and image qualities, simulating real-world scenarios in medical imaging.

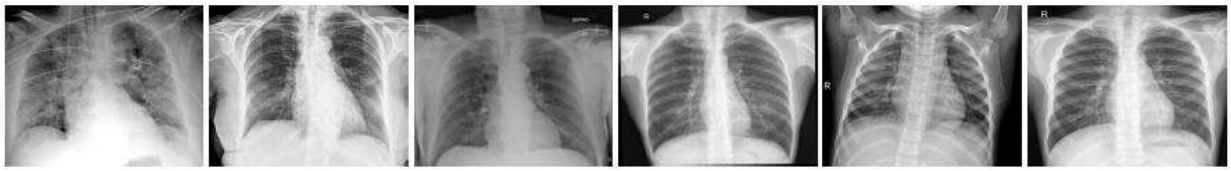


Fig 46. Images used in our experiments soured from the COVID-19 dataset, From the left to the right are the COVID-19-pneumonia-12, COVID-19-pneumonia-14-PA, COVID-19--pneumonia-35-1, NORMAL2-IM-00441-0001, NORMAL2-IM-0894-0001, and NORMAL2-IM-1275-0001 images

Used Metrics To assess the performance of the proposed watermarking method, we employed the same metrics as outlined in the first method, including Peak Signal-to-Noise Ratio (PSNR), Normalized Correlation (NC), Structural Similarity Index (SSIM), and Bit Error Rate (BER). These metrics provide a comprehensive framework for evaluating the

proposed scheme, addressing the critical aspects of image quality, watermark integrity, and robustness against various challenges.

V.3.2. Experimental results

a) Imperceptibility Assessment

We evaluated the imperceptibility of the watermarked images to the human eye using two key metrics: the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index (SSIM). Higher PSNR values indicate better quality, with minimal visual distortion caused by the watermark. The PSNR provides a more comprehensive assessment of visual fidelity, as the PSNR can be fooled by noise that appears similar to the image details. The results assessing the effect of different watermark sizes on imperceptibility are summarized in Table 1.

Table 11. Imperceptibility assessment results across different watermark sizes

Watermark Size (Bytes)	**PSNR (dB)**	**SSIM**
96	59.843 ± 1.738	0.91 ± 0.091
411	59.835 ± 1.735	0.85 ± 0.057
10779	59.554 ± 1.549	0.81 ± 0.027
43228	58.867 ± 1.122	0.80 ± 0.026
97493	57.941 ± 0.698	0.79 ± 0.021

From Table 1, it is evident that smaller watermark sizes tend to yield slightly higher PSNR values, indicating superior imperceptibility. This observation aligns with the established principle that smaller watermarks are less likely to introduce noticeable visual artifacts in the watermarked images. However, it is noteworthy that even with larger watermark sizes, the imperceptibility remains impressively high, underscoring the robustness of our watermarking technique.

For instance, when utilizing a medium watermark size of 103 bytes, the resulting image exhibits almost imperceptible changes, highlighting the method's effectiveness. Human observers typically would not perceive any significant quality difference between the original and watermarked images if the PSNR value exceeds 50 dB. This demonstrates our method’s capability to preserve the visual quality of medical images while successfully embedding the watermark.

The findings from this analysis reinforce the practical applicability of our watermarking approach in medical imaging scenarios. The ability to embed watermarks with minimal impact on image quality, even with larger sizes, enhances the protection of sensitive

medical data without compromising the diagnostic integrity of the images. To further elucidate the impact of varying watermark sizes on the imperceptibility of our watermarking technique, we provide visual illustrations in Figure 47. This figure showcases the embedding of watermark images of different sizes within cover images, ranging from small to large. As depicted, the quality and structural similarity between the watermarked images (b)–(c) and the cover image (a) remain intact, illustrating that the integrity of the images is maintained regardless of the watermark size.

To establish the superiority of imperceptibility, we conducted a comparative analysis of various watermarking methods. The primary objective was to evaluate each technique's effectiveness in concealing watermarks within images while minimizing visual distortion. Notably, our method, as illustrated in Table 12, surpassed competing approaches by achieving the highest imperceptibility score. This outcome confirms our method's exceptional ability to seamlessly integrate watermarks into medical images while preserving their original visual quality.

In conclusion, our approach demonstrates not only the capacity for effective watermark embedding but also the preservation of essential details in medical imaging, which is crucial for maintaining the diagnostic value of these images.

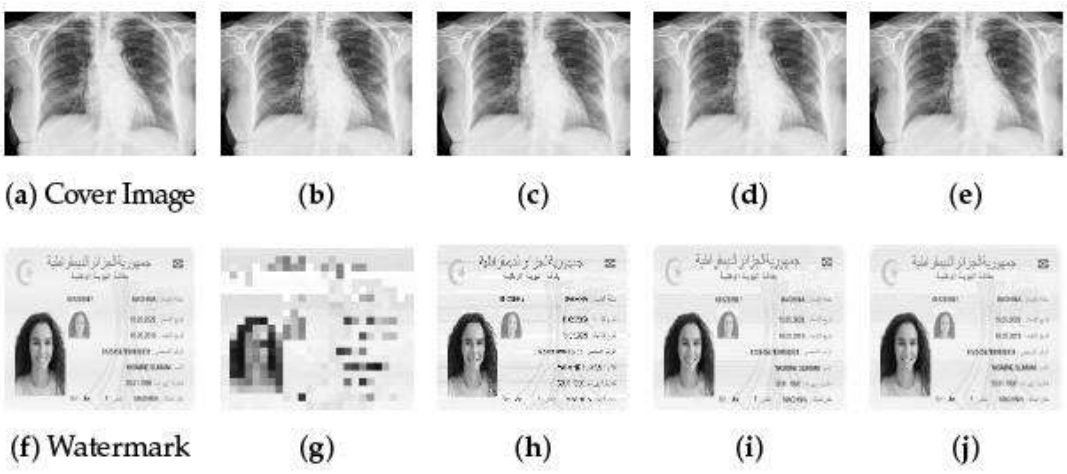


Fig 47. Illustration of watermark embedding and extraction process. (a) Cover image where the information will be hidden. (f) Watermark. (b-e) are watermarked images with increasingly larger watermark sizes. (g-j) are the corresponding extracted watermarks from the watermarked images.

Table 12. Comparaison of PSNR and SSIM scores

Method	PSNR	SSIM
Anand et al.[31]	44.19	0.85
Verma et al.[76]	41.67	0.82
Geetha et al.[68]	42.34	0.84
Gangadhar et al.[34]	45.22	0.87
Devi et al.[46]	46.64	0.89
Khaldi et al.[56]	42.97	0.83
Hurrah et al.[41]	47	0.91
Our method (Avg)	59.84	0.91

b) Robustness Evaluation

In addition to assessing imperceptibility, we also evaluated the robustness of our watermarking technique against common image processing operations and attacks. Robustness testing is critical to ensure that the embedded watermark remains intact and detectable even after various modifications to the watermarked image. Factors such as random noise or intentional tampering can significantly affect the integrity of the extracted watermark, making it essential to evaluate how well the watermark withstands these challenging conditions.

Fig. 48 illustrates the watermark extraction process both before and after applying various image attacks, including Gaussian noise and salt-and-pepper (S&P) noise. These results provide compelling evidence of the robustness of our watermarking method under different types of noise. The watermark's integrity remains largely preserved even when subjected to these common distortions, showcasing the effectiveness of our technique in maintaining watermark visibility and extractability.

The resilience of the watermark against these challenges is particularly important in medical imaging, where image quality is paramount for accurate diagnosis and treatment. By demonstrating that our method can successfully retain the watermark under adverse conditions, we reaffirm its applicability in real-world scenarios, where images are often exposed to various forms of degradation.

Overall, the ability of our watermarking technique to withstand typical image processing operations reinforces its reliability and effectiveness in safeguarding sensitive medical information. This robustness, combined with the previously noted imperceptibility,

positions our method as a strong candidate for practical applications in medical image watermarking.

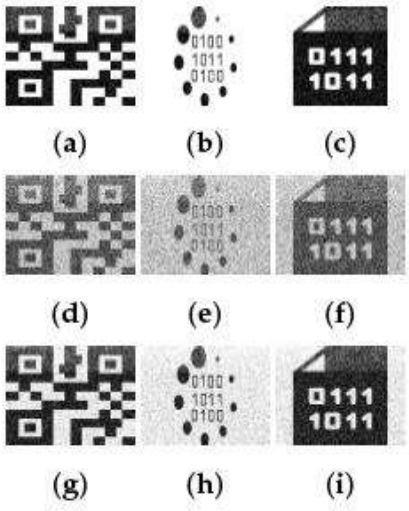


Fig 48. The effect of noise on the watermark-extration process. (a-c) Watermmarks extracted without image tamperinf. ((d-f) Extracted watermarks after applying Gaussian noise. (g-i) Extracted watermarks under salt and pepper noise.

Table 13 summarizes the results of our robustness testing, showing the performance of the watermark under each operation or attack. The resilience of the watermark against these challenges is particularly important in medical imaging, where image quality is paramount for accurate diagnosis and treatment. By demonstrating that our method can successfully retain the watermark under adverse conditions, we reaffirm its applicability in real-world scenarios, where images are often exposed to various forms of degradation.

Overall, the ability of our watermarking technique to withstand typical image processing operations reinforces its reliability and effectiveness in safeguarding sensitive medical information. This robustness, combined with the previously noted imperceptibility, positions our method as a strong candidate for practical applications in medical image watermarking.

Table 13. Comparisons of robustness (BER and NC) among different methods using a watermark of size 96 bytes (768 bits)

Attacks	[31]	[76]	[68]		[34]	[46]	[56]		[41]	Our method	
	NC	NC	NC	BER	NC	NC	NC	BER	NC	NC	BER
Median filter (kernel)	0.94 (2x2)	-	0.58 (3x3)	0.41	-	-	0.89	0.24	0.59 (3x3)	0.86 (3x3)	0.13
Salt and Pepper	0.96 (.001)	0.99	0.95 (0.01)	0.048	0.84 (0.01)	-	0.95	0.10	0.66 (0.01)	0.97 (0.01)	0.03
Gaussian Noise	0.32 (0.05)	0.62	0.50 (0.02)	0.49	0.99 (0.01)	0.99	0.93	0.21	0.48 (0.01)	0.97 (0.05)	0.01
Scaling	0.71 (x2)	-	0.47	0.47	0.74	1	0.74	0.32	0.52	0.78 (x2)	0.12
Histogram Equalization	0.87	-	-	-	0.98	1	0.89	0.23	0.54	0.8	0.08
JPEG Compression	0.98 (60)	-	0.55 (90)	0.44	-	0.99	0.91 (20)	0.13	0.54	0.98 (20)	0.08
Sharpening	0.65 (01)	-	0.50	0.49	-	0.93	0 .88	0.21	0.54	0.66 (0.1)	0.15

c) Computational Efficiency

The computation efficiency of the proposed approach, leveraging Discrete Wavelet Transform (DWT) alongside image gradient scaling using Sobel edge detection for watermark embedding in the Least Significant Bit (LSB), entails a multi-faceted process. Integrating Sobel edge detection adds a computational layer to identify image gradients efficiently, enhancing the security of watermark embedding. Considering N as the number of pixels in the image, M as the number of bits to be embedded, and L as the number of decomposition levels in the DWT, the overall complexity comprises three main components: DWT processing $O(LN \log N)$, Sobel edge detection $O(N)$, and LSB embedding $O(MN)$. Thus, the total complexity is $O(LN \log N) + O(N) + O(MN)$, with the dominant factor contingent on the interplay between L, M, and N. Knowing that $M << N$, we can say with certainty that the increase of time consumption over the size of watermark is Linear. In the subsequent Figure 44, we delineate the computational watermarking/extraction time over variations in watermark size. As anticipated from Figure 44, the consumed time exhibits a linear correlation with the size of the watermark. This linear relationship suggests that as the size of the watermark increases, the time required for computation also increases proportionally. This

stability implies that the proposed watermarking approach can efficiently handle varying sizes of watermarks, which is crucial for applications such as video processing where watermark sizes may vary.

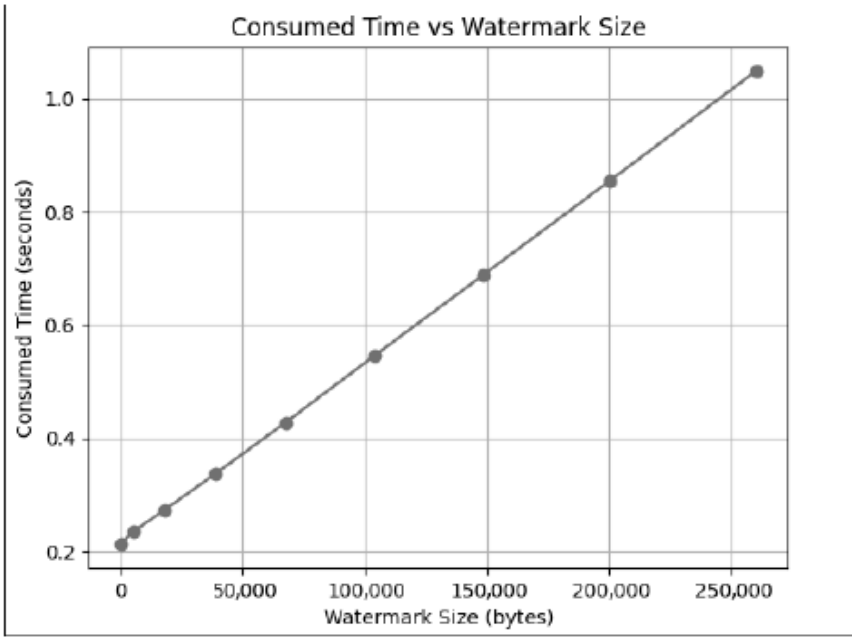


Fig 49. The impact of watermark size on the watermarking process in terms of calculation time.

. In summary, the proposed watermarking method not only achieves high levels of robustness and imperceptibility but also ensures computational efficiency. This characteristic enhances its practicality for real-world implementations, ensuring that it meets the demands of diverse applications in the medical field and beyond. Our research introduces a pioneering approach to blind and robust watermarking for medical images, leveraging the capabilities of the Discrete Wavelet Transform (DWT) combined with gradient analysis. By strategically utilizing the LL sub-band, along with gradient magnitude and gradient direction, our method strikes a crucial balance between imperceptibility and robustness. This ensures the protection of sensitive medical data while maintaining the visual integrity of the images.

V.4. Third approach: watermarking based on Deep learning feature

V.4.1. Experimental setup

In this section, we evaluate the proposed zero-watermarking method tailored for medical images, combining deep learning with watermarking techniques. The dataset utilized for this experiment is sourced from the “brain mri images” dataset, available on Kaggle (https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection,

accessed on 31 July 2024). This dataset consists of paired CT and MRI images, enabling us to assess the robustness of the watermarking method across different imaging modalities. The images are provided in JPG format..

For feature extraction, a pre-trained VGG16 model, known for its powerful feature extraction capabilities, is employed to capture essential details from the medical images. This combination of deep learning and zero-watermarking ensures that the watermark remains embedded within the feature space, rather than directly modifying the image pixels.

The proposed method was evaluated using a suite of metrics including PSNR, NC, BER. These metrics provide a comprehensive assessment of both the imperceptibility of the embedded watermark and the technique's robustness against various attacks.

To further validate the method's ability to distinguish between different images, correlation coefficients (CC) were calculated. The CC values between pairs of images were consistently found to be less than 0.5 and greater than -0.5, indicating a lack of strong positive or negative correlation. This demonstrates the method's effectiveness in preserving image uniqueness. The experimental results, depicted in Fig 50 and Table 14, confirm the proposed method's performance in terms of watermark imperceptibility and robustness. A CC value greater than 0.5 would suggest a strong positive correlation between images, while a value less than -0.5 would indicate a strong negative correlation.

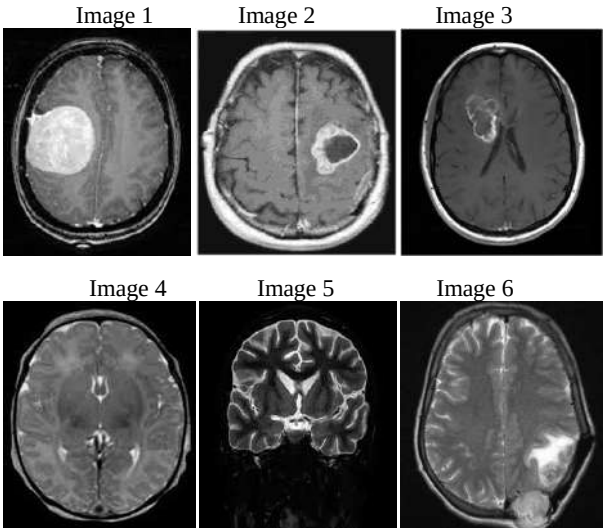


Fig 50. Six of medical images used

Table 14. Correlation coefficient between different images

	Img1	Img2	Img3	Img4	Img5	Img6	Img7	Img8
Img1	1	0.34	0.28	0.34	0.48	0.42	0.20	-0.47
Img2	0.34	1	0.30	0.35	0.31	0.25	0.13	-0.45
Img3	0.28	0.30	1	0.15	0.20	0.21	0.15	-0.36
Img4	0.34	0.35	0.15	1	0.31	0.21	0.14	-0.40
Img5	0.48	0.31	0.20	0.31	1	0.44	0.35	-0.40
Img6	0.42	0.25	0.21	0.21	0.44	1	0.40	-0.35
Img7	0.20	0.13	0.15	0.14	0.35	0.40	1	-0.23
Img8	-0.47	-0.45	-0.36	-0.40	-0.40	-0.35	-0.23	1

V.4.2. Experimental results

The proposed zero-watermarking method was rigorously tested on medical images subjected to a diverse range of attacks. The primary goal was to assess the watermark's resilience under varying attack types, intensities, and conditions. To ensure comprehensive evaluation, attacks were applied to eight distinct images from the dataset.

a) Robustness test

➢ Conventional Attacks :

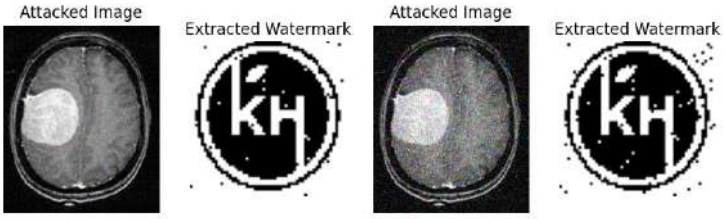
To evaluate the algorithm's robustness against non-geometric attacks, we conducted experiments with gradually increasing attack intensities. The results demonstrate the algorithm's resilience to such attacks.

▪ Gaussian Noise Attack

As illustrated in Table 11 and Fig. 46, experiments were conducted using Gaussian noise with varying attack intensities. When the Gaussian noise interference coefficient was set to 1%, the watermark achieved an NC value of 0.9988 and a BER of 0.0006. Even with a substantial increase in interference, setting the coefficient to 40 resulted in an NC of 0.9622 and a BER of 0.017, demonstrating the ability to extract relatively complete watermark information.

Table 15. PSNR, NC and BER value of image after being attacked by Gaussian noise

Gaussian Noise attack intensity	1%	5%	10%	20%	30%	40%
PSNR (dB)	47.02	34.40	28.55	22.75	19.40	17.05
NC	0.9988	0.9955	0.9886	0.9771	0.9688	0.9622
BER	0.0006	0.0022	0.0057	0.0114	0.0155	0.0017



(a) (b) (c) (d)

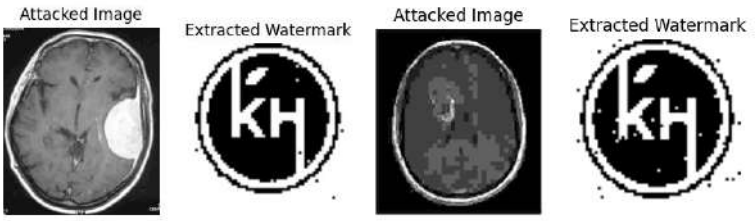
Fig 51. Image under gaussian noise attack. (a) Interference coefficient of 15%; (b) Extracted watermark with Gaussian an interference coefficient of 15%; (c) Interference coefficient of 35%; (d) Extracted watermark with Gaussian interference coefficient of 35%.

- JPEG Compression Attack

JPEG compression is a widely adopted technique in image processing. To evaluate the algorithm's resilience to JPEG attacks, experiments were conducted at various compression qualities. As shown in Table 12, even with a compression quality as low as 2, the extracted watermark maintained an NC value of 0.967. The corresponding extracted watermark image and the attacked medical image, depicted in Fig. 47, demonstrate minimal visual degradation, highlighting the algorithm's effectiveness in preserving both watermark information and image quality.

Table 16. PSNR, NC and BER value of image after compression attack

JPEG compress attack strength	2	5	10	20	30
PSNR (dB)	24.17	27.39	30.74	34.29	38.73
NC	0.9670	0.9776	0.9887	0.9934	0.9966
BER	0.0165	0.0112	0.0057	0.0033	0.0026



(a) (b) (c) (d)

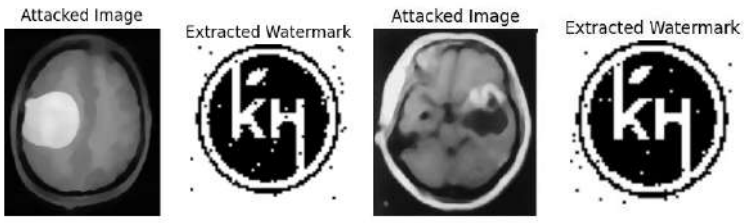
Fig 52. : Image under JPEG compression. (a) Compression quality of 10; (b) Extracted watermark with JPEG quality of 10; (c) Compression quality of 2; (d) Extracted watermark with JPEG quality of

- Median Filtering Attack

As shown in Table 13, median filter window sizes for testing are 3 × 3, 5 × 5 and 7 × 7. When filtering times is 15 times, NC values of extracted watermarks after the attack are 0.9901, 0.9725 and 0.9623 respectively. When filter window size is 7 × 7, and the filtering time is 25, the NC value is 0.9597. At this time, the valid watermark information can still be extracted, and the extracted watermark image and the attacked medical image are shown in Fig. 48.

Table 17. PSNR, NC and BER value of image after Median filtering attack

Median Filter Window Size	Filtering Times	PSNR (dB)	NC	BER
3 × 3	5	33.79	0.9912	0.0044
	15	32.84	0.9901	0.0049
	25	32.61	0.9895	0.0052
5 × 5	5	28.58	0.9779	0.0110
	15	26.54	0.9725	0.0137
	25	25.74	0.9708	0.0146
7 × 7	5	25.34	0.9687	0.0156
	15	23.24	0.9623	0.0188
	25	22.32	0.9597	0.0201



(a) (b) (c) (d)

Fig 53. Image under median filtering attack. (a) Filter median size of 5 × 5 and a filtering number of 15 times; (b) Watermark extracted when filter median size is 5 × 5 and filter times is 15 times; (c) Filter median size of 7 × 7 and filtering number of 5 times; (d) Watermark extracted when filter median size is 7 × 7 and filter times is 5 times.

- Geometric Attack

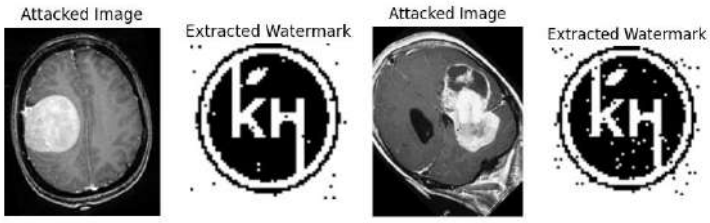
The content of this part presents experimental data on the algorithm's resilience to various geometric attacks. The results demonstrate the algorithm's exceptional ability to withstand such attacks, effectively protecting personal privacy information. Its robustness is evident in its ability to maintain watermark integrity even under significant geometric distortions.

- Rotation Attack.

After rotating the image by 35°, the NC value of the extracted watermark information is 0.9346. When the image is rotated to 80°, the NC is 0.9282. After the image is rotated by 80°, the relatively complete watermark information can still be extracted, which shows that the algorithm has good robustness. The experimental results of different degrees of rotation attacks are shown in Table 14, and the extracted images are shown in Figs. 49a and 49b.

Table 18. PSNR, NC and BER value after being attacked by Rotation attack

Rotation attack	5°	10°	15	25°	35°	45°	80°
SNR (dB)	15.31	13.38	12.37	11.31	10.81	10.58	10.26
NC	0.9668	0.9539	0.9487	0.9418	0.9346	0.9324	0.9282
BER	0.0166	0.0231	0.0256	0.0291	0.0326	0.0337	0.0358



(a) (b) (c) (d)

Fig 54. Image under Rotation attack. (a) Rotated 10° clockwise; (b) Watermark extracted after being rotated 10° clockwise; (c) Rotated 45° counterclockwise; (d) Watermark extracted after 45° counterclockwise rotation

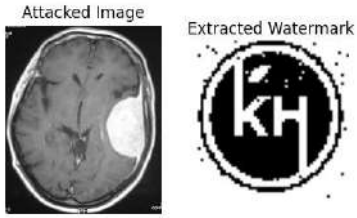
- **Scaling Attack**

To assess the algorithm's resilience to scaling attacks, experiments were conducted at various magnification levels (Table 15). Even with a 0.5x reduction, the extracted watermark maintained an NC value of 0.9915, indicating minimal impact on watermark integrity. Scaling the image up to 1.6 times resulted in an NC of 0.999, further demonstrating the algorithm's

robustness. After scaling the image 0.2 times, the extracted watermark image is clearly visible. The image is shown in Fig. 50.

Table 19. PSNR,NC and BER value of an image after scaling attack

Scaling attack factors	0.2	0.5	1	1.2	1.6	2
PSNR (dB)	-	-	-	-	-	-
NC	0.9692	0.9915	1	0.9974	0.9990	0.9966
BER	0.0154	0.0042	0	0.0013	0.0005	0.0017



(a) (b)

Fig 55. Image under Scaling attack. (a) Scaled 0.2 times; (b) Watermark extracted after scaling 0.2.

- **Translation Attack**

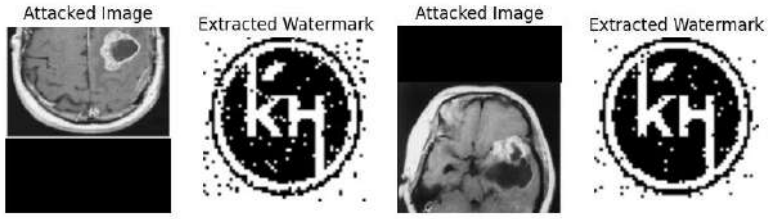
Table 20 shows the experimental data of the image after being attacked. The image is moved up by 10% and the score NC is 0.9255. When the image is moved up by 30%, the NC value of the watermark is 0.9056, close to 1.00. The medical image after 30% translation is shown in Fig. 56a, and the extracted watermark image is shown in Fig. 56b. As shown in Table 10, move the image down 15%, and the NC value is 0.95. When the image moves down 40%, the NC value of the watermark is 0.75. The medical image after 30% translation is shown in Fig. 13c, and the extracted watermark image is shown in Fig. 13d.

Table 20. PSNR,NC and BER value of image after translation attack (upward)

Translation attack (upward)	2%	5%	10%	15%	20%	30%	40%
PSNR (dB)	14.54	12.14	11.02	10.26	9.71	8.87	8.27
NC	0.9687	0.9508	0.9255	0.9115	0.9110	0.9056	0.9053
BER	0.0156	0.0246	0.0372	0.0442	0.0444	0.0471	0.0473

Table 21.PSNR,NC and BER value of image after translation attack (downward)

Translation attack (downward)	2%	5%	10%	15%	20%	30%
PSNR (dB)	14.57	12.21	11.12	10.40	9.83	8.99
NC	0.9800	0.9638	0.9534	0.9500	0.9455	0.9471
BER	0.0100	0.0181	0.0232	0.0250	0.0272	0.0264



(a) (b) (c) (d)

Fig 56. Image under Translation attack. (a) Translate 40% (up); (b) Watermark extracted after 40% translation upward; (c) Translate 30% (down); (d) Watermark extracted after 30% translation down.

The proportion of the image translating to the left is 15%. The NC is 0.9502, which is close to 1.00. When the image is translated 40% to the left, NC value is 0.9328. The image translating 40% to the left is shown in Fig. 57a, and the extracted watermark image is shown in Fig. 57b. The experimental results here are shown in Table 22.

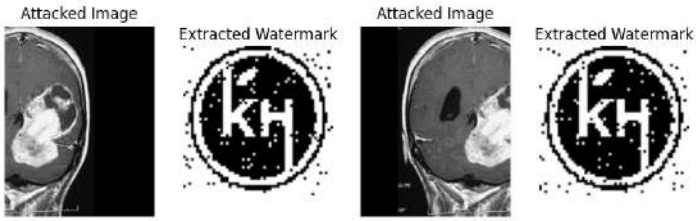
As shown in Table 12, the ratio of translation to the right of the image is 10% and the proportion of pencil NC is 0.9625. When the image is translated 40% to the right, the NC value is 0.9338. The image translated 25% to the right is shown in Fig. 57c, and the extracted watermark image is shown in Fig. 57.

Table 22. : PSNR,NC and BER value of image after translation attack (left)

left_translate_matrix (percentage)	2%	5%	10%	15%	20%	25%	40%
PSNR (dB)	14.25	11.80	10.78	10.06	9.49	8.71	8.19
NC	0.9845	0.9763	0.9612	0.9502	0.9407	0.9328	0.9328
BER	0.0077	0.0118	0.0194	0.0249	0.0296	0.0336	0.0336

Table 23. : PSNR,NC and BER value of image after translation attack (right)

right_translate_matrix (percentage)	2%	5%	10%	15%	20%	25%	40%
PSNR (dB)	14.35	11.82	10.78	10.10	9.55	9.15	8.39
NC	0.9850	0.9744	0.9625	0.9501	0.9435	0.9391	0.9338
BER	0.0075	0.0128	0.0187	0.0249	0.0282	0.0304	0.0330



(a) (b) (c) (d)

Fig 57. Image under Translation attack. (a) Translate 40% (left); (b) Translate 40% of the extracted watermark image to the left; (c) Translate 25% (right); (d) Translate 25% of the extracted watermark image to the right.

- **Cropping Attack**

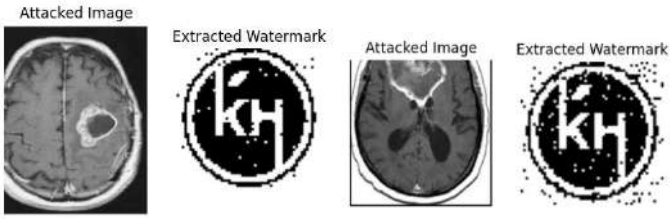
The effect of the experiment is shown in Fig. 58. Fig. 58a is the experimental object that has been cut by 30%, and Fig. 15b is the extracted watermark image. As can be seen from Table 24, when the image is cut by 15%, the NC still reaches 0.9622. When the cut ratio reaches 40%, the NC is 0.9386, which is still close to . Fig. 15c is the experimental object that was cut by 20%, and Fig. 15d is the watermark image extracted at this time. As can be seen from Table 25, the image cut ratio of 15% of the image is still 0.9111. When the cut ratio reaches 40%, the NC is 0.9157, which is still close to 1.

Table 24. PSNR,NC and BER value of image after cropping attack (X-axis)

x_cropping_matrix (percentage)	2%	5%	10%	15%	20%	25%	40%
PSNR (dB)	-	-	-	-	-	-	-
NC	0.9871	0.9811	0.9712	0.9622	0.9535	0.9462	0.9386
BER	0.0064	0.0094	0.0143	0.0189	0.0232	0.0269	0.0307

Table 25. PSNR,NC and BER value of image after cropping attack (Y axis)

y_cropping_matrix (percentage)	2%	5%	10%	15%	20%	25%	40%
PSNR (dB)	-	-	-	-	-	-	-
NC	0.9682	0.9510	0.9247	0.9127	0.9111	0.9113	0.9057
BER	0.0159	0.0244	0.0376	0.0436	0.0444	0.0443	0.0421



(a) (b) (c) (d)

Fig 58. Image under Clipping attack. (a) Cut 15%; (b) Watermark extracted after clipping attack 15%; (c) Cut 20%; (d) Watermark extracted after clipping attack (20%).

- **Algorithm Comparison**

To further illustrate the anti-geometric attack ability of the algorithm, some experimental data are compared. The comparison results are shown in Table 26. As can be seen from the table, for non geometric attacks, such as Gaussian noise and JPEG compression (the attack intensity of the two is 5%), the performance of the proposed algorithm is slightly lower than that of the algorithm proposed by others in Table 26 (all other algorithms), but the NC value of the two kinds of attacks is close to 1, which shows that the algorithm is robust. For geometric attacks, when the rotation angle reaches 10°, the NC value can reach 0.96, while the NC value of the algorithm (Yang et al. [18], Sheng et al. [64]) is 0.82 and 0.88 respectively. When the rotation angle 20° reaches 0.94, the NC values of the algorithm (Yang et al. [18], Sheng et al. [64],Gong et al. [23]) are 0.79 and 0.86 and 0.84 respectively. In this paper, it is proposed that the rotation angle of the algorithm can reach 80 degrees, which is equivalent to the rotation angle of Sheng et al. [64] algorithm, which is much higher than the rotation angle of other algorithms. For the downward translation attack with 15% attack intensity, the NC values of algorithm of Sheng et al. [64] and algorithm of Yi et al. [144] are 0.91, 0.90 respectively, and the proposed algorithm NC is 0.95. For the left translation attack with 10% attack intensity, the NC value of the algorithm Yang et al. [18] and Sheng et al. [64] are 0.63, 0.87 respectively, and the proposed algorithm NC is 0.96, also the left translation attack with 5% attack intensity, the NC value of the algorithm Gong et al. [23] is 0.90, and the proposed algorithm NC is 0.97. For the right translation attack with 5% attack intensity, the NC value of the algorithm (Zeng et al. [145], Yi et al. [144], Sheng et al. [64]) are 0.90, 0.90, 0.97 and the proposed algorithm NC is 0.97. When cutting in Y direction, when the shear ratio is 20%, the NC value of algorithm of Yang et al. [18] is 0.64, the NC value of algorithm of Zeng et al. [145] is 0.79, the NC value of algorithm of Sheng et al. [64] is 0.80, and the NC of the proposed algorithm is 0.91. at the same time, it shows that the algorithm has a stronger ability to resist geometric attacks and a better effect than the algorithm [1- 5].

To sum up, the proposed algorithm has good robustness and invisibility. The algorithm can effectively prevent information leakage and protect personal privacy information.

Table 26. Comparison of experimental results of different algorithms

Attack Type	Attack Intensity	Yang et al. [18]	Zeng et al. [145]	Yi et al. [144]	Sheng et al. [64]	Gong et al. [23]	Proposed Algorithm
Gaussian Noise	5%	0.92	0.79	0.90	0.83	0.89	0.9955
JPEG Compression	5%	-	0.79	0.90	0.62	0.95	0.9776
Rotation	10°	0.82	-	-	0.88	-	0.9539
	20°	0.79	-	-	0.86	0.84	0.9445
	80°	-	-	-	0.82	-	0.9282
Translation (Down)	15%	-	-	0.90	0.91	-	0.9500
Translation (Left)	10%	0.63	-	-	0.87	-	0.9612
	5%	-	-	-	0.92	0.90	0.9763
Translation (Right)	5%	-	0.90	0.90	0.97	-	0.9744
Cropping (Y-axis)	20%	0.64	0.79	-	0.80	-	0.9111
	30%	-	-	-	0.74	0.75	0.9089
Scaling	0.2	-	-	-	0.62	0.59	0.9692

Experimental results using the pre-trained VGG16 model on the ImageNet dataset with medical images demonstrate the effectiveness of the proposed zero-watermarking approach. By leveraging the model's learned features, the method successfully extracts hidden information from medical images without introducing visible distortions or compromising their diagnostic quality. The embedded watermark information was resilient to common image processing operations and attacks, ensuring the preservation of image integrity. These findings suggest that the proposed zero-watermarking approach is a viable solution for protecting medical images while maintaining their diagnostic value. Future research should focus on evaluating the method with a wider range of medical image datasets, comparing it to other zero-watermarking techniques, enhancing its robustness against more sophisticated attacks, and exploring potential real-world applications in medical imaging.

V.5. **Conclusion**

In this chapter, we thoroughly evaluated the performance of the proposed watermarking methods through a series of experimental tests. The results demonstrate notable differences between the three methods. In the first and second methods, the Local Binary

Pattern (LBP) and Gradient feature extraction techniques impacted the original medical images by 44.23 and 59.84 respectively, which remains within acceptable for medical imaging. On the other hand, the third method had no noticeable impact on the original images, which is crucial in medical applications to preserve image integrity.

When examining the resistance to common attacks, the first two methods showed some resilience against noise, compression, and filtering attacks, with Gaussian noise (factor 0.01), JPEG compression (factor 60/20), and median filtering (3x3 filter) offering only weak protection. However, the third method exhibited far superior resistance, handling stronger attack factors such as 40% for Gaussian noise, 5% for JPEG compression, and a 25x (3x3) median filter. Regarding geometric attacks, the first two methods were significantly affected, while the third method showed strong resistance, making it more robust against distortions like rotation, scaling, and translation. In terms of computational complexity, the deep learning-based approach took noticeably longer than the handcrafted methods, reflecting the greater computational demands of deep feature extraction.

In summary, the features extracted through deep learning show a substantial improvement over those extracted through traditional handcrafted techniques. This highlights the potential of deep learning for more robust watermarking, particularly in scenarios requiring high resistance to attacks and minimal impact on the original image.

VI. General Conclusion

Conclusion

This thesis presents a comprehensive exploration medical watermarking methods, with a particular focus on feature extraction for robust medical image watermarking. In an era where secure medical data transmission is vital for telemedicine and healthcare systems, our work contributes by proposing and evaluating multiple techniques that balance robustness, imperceptibility, and computational efficiency. The main contributions of this work include the proposal of three distinct watermarking approaches, each tailored to address the unique challenges of watermarking in medical imaging. These approaches integrate handcrafted techniques like LBP-DWT, gradient-based features, and deep learning models to achieve robust and secure watermarking.

To prove the efficiency of the proposed methods, extensive experimental evaluation was conducted. The experiments demonstrated that both the handcrafted and deep learning approaches effectively maintain the quality of medical images while embedding watermarks, particularly through the use of zero watermarking techniques. The handcrafted methods provided acceptable performance in terms of watermark embedding and resilience to basic attacks such as noise, compression, and filtering, although they showed limitations under more complex geometric attacks. Conversely, the deep learning-based approach significantly outperformed the handcrafted methods, particularly in robustness against a broader range of attacks while ensuring that the diagnostic integrity of medical images is preserved.

The thesis also highlights several key strengths of the proposed methods:

1. Robustness to Attacks: The deep learning method demonstrated exceptional resistance to noise, compression, and geometric distortions, far exceeding the capabilities of traditional handcrafted methods.
2. Medical Image Preservation: Both the handcrafted and deep learning approaches utilizing zero watermarking techniques ensure that the watermark embedding process does not alter the diagnostic quality of medical images, which is critical in clinical settings.
3. Computational Efficiency: While deep learning-based approaches offer superior robustness, the handcrafted techniques exhibit lower computational complexity, making them more suitable for resource-constrained environments.

4. Balanced Trade-offs: By comparing handcrafted and deep learning methods, the thesis identifies the need for hybrid approaches that can leverage the strengths of both techniques, combining the adaptability of deep learning with the computational efficiency of traditional methods.

In summary, this research provides valuable insights into the advantages and limitations of different watermarking techniques, proving the efficiency of the proposed approaches. The findings highlight that both handcrafted methods and deep learning techniques represent viable solutions for robust medical image watermarking while ensuring the preservation of image quality, particularly through the implementation of zero watermarking techniques. Future work may explore hybrid methods to balance robustness, efficiency, and image quality further.

References

[1] R. Szeliski, *Computer vision: algorithms and applications*. Springer Nature, 2022.

[2] N. Mittal, A. K. Pandit, M. Abouhawwash, and S. Mahajan, *Intelligent Systems and Applications in Computer Vision*. CRC Press, 2023.

[3] M. Boussif, N. Aloui, and A. Cherif, "DICOM imaging watermarking for hiding medical reports," *Med Biol Eng Comput*, vol. 58, no. 11, pp. 2905–2918, Nov. 2020, doi: 10.1007/s11517-020-02269-8.

[4] O. Aiadi and B. Khaldi, "A fast lightweight network for the discrimination of COVID-19 and pulmonary diseases," *Biomedical Signal Processing and Control*, vol. 78, p. 103925, 2022.

[5] C. O. Alenoghena *et al.*, "Telemedicine: A survey of telecommunication technologies, developments, and challenges," *Journal of Sensor and Actuator Networks*, vol. 12, no. 2, p. 20, 2023.

[6] D. K. Mahto and A. K. Singh, "A survey of color image watermarking: State-of-the-art and research directions," *Computers & Electrical Engineering*, vol. 93, p. 107255, 2021.

[7] K. Balasamy and D. Shamia, "Feature extraction-based medical image watermarking using fuzzy-based median filter," *IETE Journal of Research*, vol. 69, no. 1, pp. 83–91, 2023.

[8] X. Zhong, A. Das, F. Alrasheedi, and A. Tanvir, "A brief, in-depth survey of deep learning-based image watermarking," *Applied Sciences*, vol. 13, no. 21, p. 11852, 2023.

[9] K. M. Hosny, A. Magdi, O. ElKomy, and H. M. Hamza, "Digital image watermarking using deep learning: A survey," *Computer Science Review*, vol. 53, p. 100662, 2024.

[10] H. k.hussein, R. Muhajjar, and B. Mahdi, "Survey: Recent Techniques of Image Fragile Watermarking," *International Journal of Engineering Intelligent Systems for Electrical Engineering and Communications*, vol. 22, pp. 135–145, Jun. 2022, doi: 10.33103/uot.ijccce.22.2.12.

[11] C. Wang, H. Zhang, and X. Zhou, "LBP and DWT Based Fragile Watermarking for Image Authentication.," *Journal of Information Processing Systems*, vol. 14, no. 3, 2018.

[12] H. Liu, X. Yao, and J. Huang, "Semi-Fragile Zernike Moment-Based Image Watermarking for Authentication," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, p. 10, Feb. 2010, doi: 10.1155/2010/341856.

[13] G. Zhang *et al.*, "A Fragile Watermarking Scheme of Anti-deleting Features for 2D Vector Map," 2018, pp. 360–368. doi: 10.1007/978-3-319-73317-3_42.

[14] M. Hamidi, M. El Haziti, H. Cherifi, and M. El Hassouni, "A hybrid robust image watermarking method based on DWT-DCT and SIFT for copyright protection," *Journal of Imaging*, vol. 7, no. 10, p. 218, 2021.

[15] F. Kahlessenane, A. Khaldi, R. Kafi, and S. Euschi, "A DWT based watermarking approach for medical image protection," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 2, pp. 2931–2938, 2021.

[16] W. Liu, J. Li, J. Liu, and J. Ma, "Robust Zero Watermarking Algorithm for Medical Volume Data Based on LBP," in *Innovation in Medicine and Healthcare*, Y.-W. Chen, S. Tanaka, R. J. Howlett, and L. C. Jain, Eds., Singapore: Springer Nature Singapore, 2022, pp. 173–183.

[17] X. Xi, Y. Hua, Y. Chen, and Q. Zhu, "Zero-Watermarking for Vector Maps Combining Spatial and Frequency Domain Based on Constrained Delaunay Triangulation Network and Discrete Fourier Transform," *Entropy*, vol. 25, no. 4, p. 682, 2023.

[18] C. Yang, J. Li, U. Bhatti, J. Liu, J. Ma, and M. Huang, "Robust Zero Watermarking Algorithm for Medical Images Based on Zernike-DCT," *Security and Communication Networks*, vol. 2021, pp. 1–8, Nov. 2021, doi: 10.1155/2021/4944797.

[19] J. Liu *et al.*, "A Robust Zero-watermarking Based on SIFT-DCT for Medical Images in the Encrypted Domain," *Computers, Materials & Continua*, 2019, [Online]. Available: https://api.semanticscholar.org/CorpusID:203973164

[20] D. Mahapatra, P. Amrit, O. Singh, A. Singh, and A. Agrawal, "Autoencoder-convolutional neural network-based embedding and extraction model for image watermarking," *Journal of Electronic Imaging*, vol. 32, Sep. 2022, doi: 10.1117/1.JEI.32.2.021604.

[21] X. Zhong, P.-C. Huang, S. Mastorakis, and F. Y. Shih, "An automated and robust image watermarking scheme based on deep neural networks," *IEEE Transactions on Multimedia*, vol. 23, pp. 1951–1961, 2020.

[22] W. Zhang, J. Li, U. Bhatti, J. Liu, J. Zheng, and Y.-W. Chen, "Robust Multi-Watermarking Algorithm for Medical Images Based on GoogLeNet and Henon Map," *Computers, Materials & Continua*, vol. 75, pp. 565–586, Jan. 2023, doi: 10.32604/cmc.2023.036317.

[23] C. Gong, J. Liu, M. Gong, J. Li, U. A. Bhatti, and J. Ma, "Robust medical zero- watermarking algorithm based on Residual- DenseNet," *IET Biometrics*, vol. 11, no. 6, pp. 547–556, 2022.

[24] M. Jiang and H. Yang, "Reversible Multipurpose Watermarking Algorithm Using ResNet and Perceptual Hashing," *Journal of Information Processing Systems*, vol. 19, no. 6, pp. 756–766, 2023.

[25] S. A. Nawaz, J. Li, U. A. Bhatti, M. U. Shoukat, D. Li, and M. A. Raza, "Hybrid watermarking algorithm for medical images based on digital transformation and MobileNetV2," *Information Sciences*, vol. 653, p. 119810, 2024.

[26] S. A. Nawaz, J. Li, M. U. Shoukat, U. A. Bhatti, and M. A. Raza, "Hybrid medical image zero watermarking via discrete wavelet transform-ResNet101 and discrete cosine transform," *Computers and Electrical Engineering*, vol. 112, p. 108985, 2023.

[27] W. Cui *et al.*, "A Robust Zero Watermarking Algorithm for Medical Images Based on Tetrolet-DCT," in *International Conference on Cryptography and Security Systems*, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:236150844

[28] F. Sabbane and H. Tairi, "Medical image watermarking technique based on polynomial decomposition," *Multimedia Tools and Applications*, vol. 78, pp. 34129–34155, Dec. 2019, doi: 10.1007/s11042-019-08134-7.

[29] N. Rodríguez-Ortega, "Techno-Concepts for the Cultural Field: n-Dimensional Space and Its Conceptual Constellation," *Multimodal Technologies and Interaction*, vol. 6, no. 11, p. 96, 2022.

[30] B. Khaldi, O. Aiadi, and M. L. Kherfi, "Combining colour and grey level co occurrence matrix features: a comparative study," *IET Image Processing*, vol. 13, no. 9, pp. 1401–1410, 2019.

[31] A. Anand and A. K. Singh, "An improved DWT-SVD domain watermarking for medical information security," *Computer Communications*, vol. 152, pp. 72–80, Feb. 2020, doi: 10.1016/j.comcom.2020.01.038.

[32] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, Nov. 1991, doi: 10.1007/BF00130487.

[33] "A fuzzy based ROI selection for encryption and watermarking in medical image using DWT and SVD - ProQuest." Accessed: Jul. 29, 2023. [Online]. Available: https://www.proquest.com/openview/07198003d5d669b4f3b8529ae8abe260/1?pq-origsite=gscholar&cbl=54626

[34] Y. Gangadhar, V. G. Akula, and P. C. Reddy, "An evolutionary programming approach for securing medical images using watermarking scheme in invariant discrete wavelet transformation," *Biomedical Signal Processing and Control*, vol. 43, pp. 31–40, 2018.

[35] S. A. Parah, J. A. Sheikh, F. Ahad, N. A. Loan, and G. M. Bhat, "Information hiding in medical images: a robust medical image watermarking system for E-healthcare," *Multimedia Tools and Applications*, vol. 76, pp. 10599–10633, 2017.

[36] 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, Jul. 2022, doi: 10.1007/s11042-022-11956-7.

[37] B. Khaldi, O. Aiadi, and K. M. Lamine, "Image representation using complete multi-texton histogram," *Multimedia Tools and Applications*, vol. 79, no. 11–12, pp. 8267–8285, 2020.

[38] A. C. Gjelsteen *et al.*, "CT, MRI, PET, PET/CT, and Ultrasound in the Evaluation of Obstetric and Gynecologic Patients," *Surgical Clinics of North America*, vol. 88, no. 2, pp. 361–390, Apr. 2008, doi: 10.1016/j.suc.2008.01.005.

[39] S. Sharif, R. A. Naqvi, and M. Biswas, "Learning medical image denoising with deep dynamic residual attention network," *Mathematics*, vol. 8, no. 12, p. 2192, 2020.

[40] Y.-W. Chow, W. Susilo, J. Baek, and J. Kim, "QR code watermarking for digital images," presented at the Information Security Applications: 20th International Conference, WISA 2019, Jeju Island, South Korea, August 21–24, 2019, Revised Selected Papers 20, Springer, 2020, pp. 25–37.

[41] 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.

[42] D. Mata-Mendoza, M. Cedillo-Hernandez, F. Garcia-Ugalde, A. Cedillo-Hernandez, M. Nakano-Miyatake, and H. Perez-Meana, "Secured telemedicine of medical imaging based on dual robust watermarking," *The Visual Computer*, vol. 38, no. 6, pp. 2073–2090, 2022.

[43] S. Borra and R. Thanki, "A FRT-SVD based blind medical watermarking technique for telemedicine applications," *International Journal of Digital Crime and Forensics (IJDCF)*, vol. 11, no. 2, pp. 13–33, 2019.

[44] F. Liu, "A Novel Fragile Zero-watermarking Algorithm Based on Quaternion and LBP under the Background of Blockchain," in *2023 5th International Conference on Communications, Information System and Computer Engineering (CISCE)*, Apr. 2023, pp. 411–414. doi: 10.1109/CISCE58541.2023.10142515.

[45] M. AlShaikh, M. Alzaqebah, and S. Jawarneh, "Robust watermarking based on modified Pigeon algorithm in DCT domain," *Multimedia Tools and Applications*, vol. 82, no. 2, pp. 3033–3053, 2023.

[46] H. S. Devi and H. Mohapatra, "A novel robust blind medical image watermarking using rank-based DWT," *International Journal of Information Technology*, vol. 15, no. 4, pp. 1901–1909, Apr. 2023, doi: 10.1007/s41870-023-01234-6.

[47] A. Benoraira, K. Benmahammed, and N. Boucenna, "Blind image watermarking technique based on differential embedding in DWT and DCT domains," *EURASIP Journal on Advances in Signal Processing*, vol. 2015, no. 1, pp. 1–11, 2015.

[48] H. S. Devi and K. M. Singh, "Red-cyan anaglyph image watermarking using DWT, Hadamard transform and singular value decomposition for copyright protection," *Journal of Information Security and Applications*, vol. 50, p. 102424, Feb. 2020, doi: 10.1016/j.jisa.2019.102424.

[49] P. Pal, B. Jana, and J. Bhaumik, "A secure reversible color image watermarking scheme based on LBP, lagrange interpolation polynomial and weighted matrix," *Multimedia Tools and Applications*, vol. 80, no. 14, pp. 21651–21678, Jun. 2021, doi: 10.1007/s11042-021-10651-3.

[50] "A Novel Robust Reversible Watermarking Scheme for Protecting Authenticity and Integrity of Medical Images | IEEE Journals & Magazine | IEEE Xplore." Accessed: Jul. 29, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8733801

[51] F. Kahlessenane, A. Khaldi, R. Kafi, and S. Euschi, "A robust blind medical image watermarking approach for telemedicine applications," *Cluster computing*, vol. 24, no. 3, pp. 2069–2082, 2021.

[52] X. Kang, J. Huang, and Y. Q. Shi, “An Image Watermarking Algorithm Robust to Geometric Distortion,” in *Digital Watermarking*, H. J. Kim, Ed., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2003, pp. 212–223. doi: 10.1007/3-540-36617-2_18.

[53] M. Ghadi, L. Laouamer, L. Nana, and A. Pascu, “A novel zero-watermarking approach of medical images based on Jacobian matrix model,” *Security and communication networks*, vol. 9, no. 18, pp. 5203–5218, 2016.

[54] Z. B. Faheem *et al.*, “Image watermarking scheme using LSB and image gradient,” *Applied Sciences*, vol. 12, no. 9, p. 4202, 2022.

[55] H. S. Devi and H. Mohapatra, “A novel robust blind medical image watermarking using GWO optimized DWT-DCT-SVD,” *Multimedia Tools and Applications*, Apr. 2023, doi: 10.1007/s11042-023-15158-7.

[56] K. Amine, K. Redouane, and M. Bilel, “A redundant wavelet based medical image watermarking scheme for secure transmission in telemedicine applications,” *Multimedia Tools and Applications*, vol. 82, no. 5, pp. 7901–7915, 2023.

[57] S. B. B. Ahmadi, G. Zhang, and S. Wei, “Robust and hybrid SVD-based image watermarking schemes:,” *Multimedia Tools and Applications*, vol. 79, no. 1, pp. 1075–1117, Jan. 2020, doi: 10.1007/s11042-019-08197-6.

[58] H. Chen, B. D. Rouhani, X. Fan, O. C. Kilinc, and F. Koushanfar, “Performance comparison of contemporary DNN watermarking techniques,” *arXiv preprint arXiv:1811.03713*, 2018.

[59] T. K. Araghi and D. Megías, “Analysis and effectiveness of deeper levels of SVD on performance of hybrid DWT and SVD watermarking,” *Multimedia Tools and Applications*, vol. 83, no. 2, pp. 3895–3916, 2024.

[60] Z. Jia, H. Fang, and W. Zhang, “Mbrs: Enhancing robustness of dnn-based watermarking by mini-batch of real and simulated jpeg compression,” presented at the Proceedings of the 29th ACM international conference on multimedia, 2021, pp. 41–49.

[61] D. Li, L. Deng, B. B. Gupta, H. Wang, and C. Choi, “A novel CNN based security guaranteed image watermarking generation scenario for smart city applications,” *Information Sciences*, vol. 479, pp. 432–447, 2019.

[62] T. Huang, J. Xu, S. Tu, and B. Han, “Robust zero-watermarking scheme based on a depthwise overparameterized VGG network in healthcare information security,” *Biomedical Signal Processing and Control*, vol. 81, p. 104478, 2023.

[63] S. Boujerfaoui, R. Riad, H. Douzi, F. Ros, and R. Harba, “Image watermarking between conventional and learning-based techniques: a literature review,” *Electronics*, vol. 12, no. 1, p. 74, 2022.

[64] M. Sheng, J. Li, U. A. Bhatti, J. Liu, M. Huang, and Y. W. Chen, “Zero watermarking algorithm for medical image based on Resnet50-DCT,” *CMC-Computers Materials & Continua*, vol. 75, no. 1, pp. 293–309, 2023.

[65] Q. Wei, H. Wang, and G. Zhang, “A robust image watermarking approach using cycle variational autoencoder,” *Security and Communication Networks*, vol. 2020, no. 1, p. 8869096, 2020.

[66] K. Kishan and B. V. Kumar, “Efficient large invisible color watermark embedding using conditional deep autoencoder model for medical applications,” *Measurement: Sensors*, vol. 29, p. 100850, 2023.

[67] F. Kahlessenane, A. Khaldi, M. R. Kafi, and S. Euschi, “A color value differentiation scheme for blind digital image watermarking,” *Multimed Tools Appl*, vol. 80, no. 13, pp. 19827–19844, May 2021, doi: 10.1007/s11042-021-10713-6.

[68] 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. 19–20, pp. 12869–12890, 2020.

[69] S. Jaiswal and M. K. Pandey, “Robust Image Watermarking Using Arnold Map and Adaptive Threshold Value in LWT Domain,” in *Machine Intelligence Techniques for Data Analysis and Signal Processing*, D. S. Sisodia, L. Garg, R. B. Pachori, and M. Tanveer, Eds., Singapore: Springer Nature Singapore, 2023, pp. 343–355.

[70] N. Sharma, A. Anand, and A. K. Singh, “Bio-signal data sharing security through watermarking: a technical survey,” *Computing*, vol. 103, no. 9, pp. 1883–1917, 2021.

[71] E. Elbasi and V. Kaya, “Robust medical image watermarking using frequency domain and least significant bits algorithms,” presented at the 2018 International Conference on Computing Sciences and Engineering (ICCSE), IEEE, 2018, pp. 1–5.

[72] S. Liu, Z. Pan, and H. Song, “Digital image watermarking method based on DCT and fractal encoding,” *IET Image Process.*, vol. 11, pp. 815–821, 2017, [Online]. Available: https://api.semanticscholar.org/CorpusID:23917100

[73] A. Zear and A. Singh, “A proposed secure multiple watermarking technique based on DWT, DCT and SVD for application in medicine,” *Multimedia Tools and Applications*, vol. 77, Feb. 2018, doi: 10.1007/s11042-016-3862-8.

[74] R. Agrawal, M. Sharma, and B. Singh, “Hiding Patient Information in Medical Images: A Robust Watermarking Algorithm for Healthcare System,” 2021, pp. 245–261. doi: 10.1007/978-981-15-6329-4_22.

[75] A. Singh, M. Dave, and A. Mohan, “Hybrid technique for robust and imperceptible multiple watermarking using medical images,” *Multimedia Tools and Applications*, vol. 75, Jul. 2015, doi: 10.1007/s11042-015-2754-7.

[76] U. Verma and N. Sharma, “Hybrid mode of medical image watermarking to enhance robustness and imperceptibility,” *Int J Innov Technol Explor Eng*, vol. 9, no. 1, pp. 351–359, 2019.

[77] M. Zairi, T. Boujiha, and A. Ouelli, “Secure fragile watermarking based on Huffman encoding and optimal embedding strategy,” *Indon. J. Electr. Eng. Comput. Sci*, vol. 29, no. 2, pp. 1132–1139, 2023.

[78] S. Gull, N. Loan, S. Parah, J. Sheikh, and G. Bhat, “An efficient watermarking technique for tamper detection and localization of medical images,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, May 2020, doi: 10.1007/s12652-018-1158-8.

[79] M. Cedillo-Hernandez, A. Cedillo-Hernandez, M. Nakano-Miyatake, and H. Perez-Meana, “Improving the management of medical imaging by using robust and secure dual watermarking,” *Biomedical Signal Processing and Control*, vol. 56, p. 101695, Feb. 2020, doi: 10.1016/j.bspc.2019.101695.

[80] M. Jana, D.-B. Jana, and S. Joardar, “Local feature based self-embedding fragile watermarking scheme for tampered detection and recovery utilizing AMBTC with fuzzy logic,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, Jan. 2022, doi: 10.1016/j.jksuci.2021.12.011.

[81] M. Sayah, K. Mohamed Redouane, and A. Khaldi, “A wavelet based medical image watermarking scheme for secure transmission in telemedicine applications,” *Microprocessors and Microsystems*, vol. 90, p. 104490, Feb. 2022, doi: 10.1016/j.micpro.2022.104490.

[82] S. Singh, R. Singh, A. K. Singh, and T. J. Siddiqui, “SVD-DCT Based Medical Image Watermarking in NSCT Domain,” 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:125456491

[83] H. Wu, J. Huang, and Y. Q. Shi, “A Reversible Data Hiding Method with Contrast Enhancement for Medical Images,” *Journal of Visual Communication and Image Representation*, vol. 31, pp. 146–153, Jun. 2015, doi: 10.1016/j.jvcir.2015.06.010.

[84] K. Swaraja, “Medical image region based watermarking for secured telemedicine,” *Multimedia Tools and Applications*, vol. 77, pp. 1–32, Nov. 2018, doi: 10.1007/s11042-018-6020-7.

[85] R. E. Arevalo-Ancona and M. Cedillo-Hernández, “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*, 2023, [Online]. Available: https://api.semanticscholar.org/CorpusID:259311840

[86] Y. Luo, J. Li, U. A. Bhatti, M. Huang, F. Dong, and Y. Li, “Robust Zero-Watermarking Algorithm for Medical Images Based on Direction Gradient Histogram and DCT,” presented at the 2023 IEEE 6th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), IEEE, 2023, pp. 575–583.

[87] Y. X. Fang *et al.*, “Robust zero-watermarking algorithm for medical images based on SIFT and Bandelet-DCT,” *Multimedia Tools and Applications*, vol. 81, pp. 16863–16879, 2022, [Online]. Available: https://api.semanticscholar.org/CorpusID:247263753

[88] G. Sun, J. Li, U. A. Bhatti, J. Ma, F. Dong, and Y. Li, “Robust Zero-Watermarking Algorithm for Medical Images Based on AGAST-LATCH and DCT,” *2023 26th ACIS International Winter Conference on Software Engineering, Artificial Intelligence,*

Networking and Parallel/Distributed Computing (SNPD-Winter), pp. 298–305, 2023, [Online]. Available: https://api.semanticscholar.org/CorpusID:261328626

[89] V. Seenivasagam and R. Velumani, "A QR Code Based Zero-Watermarking Scheme for Authentication of Medical Images in Teleradiology Cloud," *Computational and Mathematical Methods in Medicine*, vol. 2013, 2013, [Online]. Available: https://api.semanticscholar.org/CorpusID:9797354

[90] M. Magdy, N. I. Ghali, S. Ghoniemy, and K. M. Hosny, "Multiple Zero-Watermarking of Medical Images for Internet of Medical Things," *IEEE Access*, vol. PP, pp. 1–1, 2022, [Online]. Available: https://api.semanticscholar.org/CorpusID:248061673

[91] J. Zhu, "HiDDeN: hiding data with deep networks," *arXiv preprint arXiv:1807.09937*, 2018.

[92] H. K. Singh and A. K. Singh, "Digital image watermarking using deep learning," *Multimedia Tools and Applications*, vol. 83, no. 1, pp. 2979–2994, 2024.

[93] C. Zhang, P. Benz, A. Karjauv, G. Sun, and I. S. Kweon, "UDH: Universal Deep Hiding for Steganography, Watermarking, and Light Field Messaging," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, Eds., Curran Associates, Inc., 2020, pp. 10223–10234. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2020/file/73d02e4344f71a0b0d51a925246990e7-Paper.pdf

[94] R. Sinhal and I. A. Ansari, "Machine learning based multipurpose medical image watermarking," *Neural Computing and Applications*, vol. 35, no. 31, pp. 23041–23062, 2023.

[95] S. Jaiswal and M. K. Pandey, "Color Watermark Extraction Using Deep Neural Network in IWT Domain with PCA-Based Statistical Feature Reduction," *SN Computer Science*, vol. 4, no. 5, p. 669, 2023.

[96] M. Rai, S. Goyal, and M. Pawar, "An optimized deep fusion convolutional neural network-based digital color image watermarking scheme for copyright protection," *Circuits, Systems, and Signal Processing*, vol. 42, no. 7, pp. 4019–4050, 2023.

[97] Y. Liu, M. Guo, J. Zhang, Y. Zhu, and X. Xie, "A novel two-stage separable deep learning framework for practical blind watermarking," presented at the Proceedings of the 27th ACM International conference on multimedia, 2019, pp. 1509–1517.

[98] C. Zhang, A. Karjauv, P. Benz, and I. S. Kweon, "Towards robust deep hiding under non-differentiable distortions for practical blind watermarking," presented at the Proceedings of the 29th ACM international conference on multimedia, 2021, pp. 5158–5166.

[99] B. Chen, Y. Wu, G. Coatrieux, X. Chen, and Y. Zheng, "JSNet: a simulation network of JPEG lossy compression and restoration for robust image watermarking against JPEG attack," *Computer Vision and Image Understanding*, vol. 197, p. 103015, 2020.

[100] H. Fang, Z. Jia, Z. Ma, E.-C. Chang, and W. Zhang, "PIMoG: An Effective Screen-shooting Noise-Layer Simulation for Deep-Learning-Based Watermarking Network," in

Proceedings of the 30th ACM International Conference on Multimedia, in MM '22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 2267–2275. doi: 10.1145/3503161.3548049.

[101] W. Gu, C.-C. Chang, Y. Bai, Y. Fan, L. Tao, and L. Li, "Anti-screenshot watermarking algorithm for archival image based on deep learning model," *Entropy*, vol. 25, no. 2, p. 288, 2023.

[102] J. Huang, T. Luo, L. Li, G. Yang, H. Xu, and C.-C. Chang, "ARWGAN: Attention-guided robust image watermarking model based on GAN," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–17, 2023.

[103] X. Luo, R. Zhan, H. Chang, F. Yang, and P. Milanfar, *Distortion Agnostic Deep Watermarking*. 2020, p. 13554. doi: 10.1109/CVPR42600.2020.01356.

[104] M. Ahmadi, A. Norouzi, N. Karimi, S. Samavi, and A. Emami, "ReDMark: Framework for residual diffusion watermarking based on deep networks," *Expert Systems with Applications*, vol. 146, p. 113157, 2020.

[105] M. Plata and P. Syga, *Robust Spatial-Spread Deep Neural Image Watermarking*. 2020, p. 70. doi: 10.1109/TrustCom50675.2020.00022.

[106] L. Zhang, W. Li, and H. Ye, "A blind watermarking system based on deep learning model," in *2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, Oct. 2021, pp. 1208–1213. doi: 10.1109/TrustCom53373.2021.00164.

[107] K. Hao, G. Feng, and X. Zhang, "Robust image watermarking based on generative adversarial network," *China Communications*, vol. 17, no. 11, pp. 131–140, 2020.

[108] H.-B. Xu, R. Wang, J. Wei, and S.-P. Lu, *A Compact Neural Network-based Algorithm for Robust Image Watermarking*. 2021. doi: 10.48550/arXiv.2112.13491.

[109] D. Mahapatra, P. Amrit, O. Singh, A. Singh, and A. Agrawal, "Autoencoder-convolutional neural network-based embedding and extraction model for image watermarking," *Journal of Electronic Imaging*, vol. 32, Sep. 2022, doi: 10.1117/1.JEI.32.2.021604.

[110] Y. Zhao, C. Wang, X. Zhou, and Z. Qin, "DARI-Mark: Deep learning and attention network for robust image watermarking," *Mathematics*, vol. 11, no. 1, p. 209, 2022.

[111] Q. Ying, H. Zhou, X. Zeng, H. Xu, Z. Qian, and X. Zhang, *Hiding Images Into Images with Real-World Robustness*. 2022, p. 115. doi: 10.1109/ICIP46576.2022.9897931.

[112] H. Fang, Z. Jia, Y. Qiu, J. Zhang, W. Zhang, and E.-C. Chang, "De-END: Decoder-driven Watermarking Network," *IEEE Transactions on Multimedia*, vol. PP, pp. 1–11, Jan. 2022, doi: 10.1109/TMM.2022.3223559.

[113] S.-M. Mun, S.-H. Nam, H. Jang, D. Kim, and H.-K. Lee, "Finding robust domain from attacks: A learning framework for blind watermarking," *Neurocomputing*, vol. 337, pp. 191–202, 2019.

[114] L. Zhu, X. Wen, L. Mo, J. Ma, and D. Wang, “Robust location-secured high-definition image watermarking based on key-point detection and deep learning,” *Optik*, vol. 248, p. 168194, 2021.

[115] A. Das and X. Zhong, *A Deep Learning-based Audio-in-Image Watermarking Scheme*. 2021, p. 5. doi: 10.1109/VCIP53242.2021.9675375.

[116] S. Ge, Z. Xia, J. Fei, Y. Tong, J. Weng, and M. Li, “A robust document image watermarking scheme using deep neural network,” *Multimedia Tools and Applications*, vol. 82, no. 25, pp. 38589–38612, 2023.

[117] X. Liao, J. Peng, and Y. Cao, “GIFMarking: The robust watermarking for animated GIF based deep learning,” *Journal of Visual Communication and Image Representation*, vol. 79, p. 103244, 2021.

[118] G. Gao, T. Xu, and F. Hua, *Robust Image Watermarking Based on Generative Adversarial Networks for Copyright Protection*. 2024. doi: 10.21203/rs.3.rs-4039149/v1.

[119] A. Fierro, M. Nakano-Miyatake, M. Cedillo-Hernandez, L. Cleofas, and H. Perez-Meana, *A Robust Image Zero-watermarking using Convolutional Neural Networks*. 2019, p. 5. doi: 10.1109/IWBF.2019.8739245.

[120] L. He, Z. He, T. Luo, and Y. Song, “Shrinkage and redundant feature elimination network-based robust image zero-watermarking,” *Symmetry*, vol. 15, no. 5, p. 964, 2023.

[121] B. Han, H. Wang, D. Qiao, J. Xu, and T. Yan, “Application of zero-watermarking scheme based on swin transformer for securing the metaverse healthcare data,” *IEEE Journal of Biomedical and Health Informatics*, 2023.

[122] V. Vukotić, V. Chappelier, and T. Furon, “Are classification deep neural networks good for blind image watermarking?,” *Entropy*, vol. 22, no. 2, p. 198, 2020.

[123] F. Haimour, R. Al-Sayyed, W. Mahafza, and O. S. Al-Kadi, “Bidirectional brain image translation using transfer learning from generic pre-trained models,” *Computer Vision and Image Understanding*, vol. 248, p. 104100, Nov. 2024, doi: 10.1016/j.cviu.2024.104100.

[124] P. Fernandez, A. Sablayrolles, T. Furon, H. Jégou, and M. Douze, *Watermarking Images in Self-Supervised Latent Spaces*. 2022, p. 3058. doi: 10.1109/ICASSP43922.2022.9746058.

[125] A. Chinnamuthu, N. Iruthayanathan, N. Kathamuthu, R. MANIKANDAN, and A. Gandomi, “Image Watermarking Based Data Hiding by Discrete Wavelet Transform Quantization Model with Convolutional Generative Adversarial Architectures,” *Applied Sciences*, vol. 13, Jan. 2023, doi: 10.3390/app13020804.

[126] H. Kandi, D. Mishra, and S. R. K. S. Gorthi, “Exploring the learning capabilities of convolutional neural networks for robust image watermarking,” *Computers & Security*, vol. 65, pp. 247–268, Mar. 2017, doi: 10.1016/j.cose.2016.11.016.

[127] A. Ferdowsi and W. Saad, *Deep Learning-Based Dynamic Watermarking for Secure Signal Authentication in the Internet of Things*. 2018, p. 6. doi: 10.1109/ICC.2018.8422728.

[128] H. K. Singh, N. Baranwal, K. N. Singh, A. K. Singh, and H. Zhou, "GAN-based watermarking for encrypted images in healthcare scenarios," *Neurocomputing*, vol. 560, p. 126853, 2023.

[129] S. Mellimi, V. Rajput, I. A. Ansari, and C. W. Ahn, "A fast and efficient image watermarking scheme based on deep neural network," *Pattern Recognition Letters*, vol. 151, pp. 222–228, 2021.

[130] A. Chacko and S. Chacko, "Deep learning-based robust medical image watermarking exploiting DCT and Harris hawks optimization," *International Journal of Intelligent Systems*, vol. 37, Nov. 2021, doi: 10.1002/int.22742.

[131] H. Fang *et al.*, "Deep template-based watermarking," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 4, pp. 1436–1451, 2020.

[132] W.-H. Kim, J. Kang, S.-M. Mun, and J.-U. Hou, "Convolutional neural network architecture for recovering watermark synchronization," *Sensors*, vol. 20, no. 18, p. 5427, 2020.

[133] Y.-P. Chen, T.-Y. Fan, and H.-C. Chao, "Wmnet: A lossless watermarking technique using deep learning for medical image authentication," *Electronics*, vol. 10, no. 8, p. 932, 2021.

[134] D. K. Mahto, A. Anand, and A. K. Singh, "Hybrid optimisation-based robust watermarking using denoising convolutional neural network," *Soft Computing*, vol. 26, no. 16, pp. 8105–8116, 2022.

[135] G. Liu, R. Xiang, J. Liu, R. Pan, and Z. Zhang, "An invisible and robust watermarking scheme using convolutional neural networks," *Expert Systems with Applications*, vol. 210, p. 118529, 2022.

[136] M. M. Darwish, A. A. Farhat, and T. M. El-Gindy, "Convolutional neural network and 2D logistic-adjusted-Chebyshev-based zero-watermarking of color images," *Multimedia Tools and Applications*, vol. 83, no. 10, pp. 29969–29995, 2024.

[137] R. Xiang, G. Liu, K. Li, J. Liu, Z. Zhang, and M. Dang, "Zero-watermark scheme for medical image protection based on style feature and ResNet," *Biomedical Signal Processing and Control*, vol. 86, p. 105127, 2023.

[138] D. Li *et al.*, "Hybrid encrypted watermarking algorithm for medical images based on DCT and improved DarkNet53," *Electronics*, vol. 12, no. 7, p. 1554, 2023.

[139] A. Anand, J. Bedi, A. Aggarwal, M. A. Khan, and I. Rida, "Authenticating and securing healthcare records: A deep learning-based zero watermarking approach," *Image and Vision Computing*, vol. 145, p. 104975, 2024.

[140] F. Dong, J. Li, U. A. Bhatti, J. Liu, Y.-W. Chen, and D. Li, “Robust zero watermarking algorithm for medical images based on improved NasNet-mobile and DCT,” *Electronics*, vol. 12, no. 16, p. 3444, 2023.

[141] B. Li and J. Xu, “Period of Arnold transformation and its application in image scrambling,” *Journal of Central South University of Technology*, vol. 12, no. Suppl 1, pp. 278–282, 2005.

[142] S. P. Vaidya, “Fingerprint-based robust medical image watermarking in hybrid transform,” *The Visual Computer*, vol. 39, no. 6, pp. 2245–2260, Jun. 2023, doi: 10.1007/s00371-022-02406-4.

[143] M. Chhabra and R. Kumar, “An Advanced VGG16 Architecture-Based Deep Learning Model to Detect Pneumonia from Medical Images,” in *Emergent Converging Technologies and Biomedical Systems*, N. Marriwala, C. C. Tripathi, S. Jain, and S. Mathapathi, Eds., Singapore: Springer Singapore, 2022, pp. 457–471.

[144] D. Yi *et al.*, “A robust zero-watermarkinging algorithm based on PHTs-DCT for medical images in the encrypted domain,” presented at the Innovation in Medicine and Healthcare: Proceedings of 9th KES-InMed 2021, Springer, 2021, pp. 101–113.

[145] C. Zeng *et al.*, “Multi-watermarking algorithm for medical image based on KAZE-DCT,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, no. 2, pp. 1735–1743, 2024.