electronics

MDPI

Review

Image Watermarking between Conventional and Learning-Based Techniques: A Literature Review

Said Boujerfaoui [1,*], Rabia Riad [2], Hassan Douzi [1], Frédéric Ros [3] and Rachid Harba [3]

[1] IRF-SIC, University of Ibn Zohr, Agadir 80000, Morocco
[2] PETI, University of Ibn Zohr, Ouarzazate 45000, Morocco
[3] PRISME, University of Orléans, 45000 Orléans, France
* Correspondence: said.boujerfaoui@edu.uiz.ac.ma

Abstract: Currently, most transactions and exchanges are conducted through the Internet thanks to technological tools, running the risk of the falsification and distortion of information. This is due to the massive demand for the virtual world and its easy access to anyone. Image watermarking has recently emerged as one of the most important areas for protecting content and enhancing durability and resistance to these kinds of attacks. However, there is currently no integrated technology able to repel all possible kinds of attacks; the main objective of each technology remains limited to specific types of applications, meaning there are multiple opportunities to contribute to the development of this field. Recently, the image watermarking field has gained significant benefits from the sudden popularity of deep learning and its outstanding success in the field of information security. Thus, in this article, we will describe the bridge by which the watermarking field has evolved from traditional technology to intelligent technologies based on deep learning.

Keywords: image watermarking; deep learning; digital images; copyright protection; information security; visual imperceptibility; robustness; review

Citation: Boujerfaoui, S.; Riad, R.; Douzi, H.; Ros, F.; Harba, R. Image Watermarking between Conventional and Learning-Based Techniques: A Literature Review. *Electronics* **2023**, *12*, 74. https://doi.org/10.3390/electronics12010074

Academic Editor: Chiman Kwan

Received: 3 November 2022
Revised: 17 December 2022
Accepted: 19 December 2022
Published: 25 December 2022

1. Introduction

Digital media in this era has become an integral component of every individual's daily life. Large amounts of data are stored and exchanged easily at an incredible speed [1] due to the current technological advancements that provide environments for data sharing and a vast amount of storage space. However, there are many concerns about the illegal piracy of copyrighted media. This issue gave rise to the concept of digital watermarking. Since the 1990s, it has become a new research direction and is attracting significant interest in the scientific community. Watermarking is considered a relatively young discipline and appears to be a suitably efficient solution for different security fields, such as copyright protection and authentication [2]. Figure 1 shows the chronological development of embedding operations in watermarking.

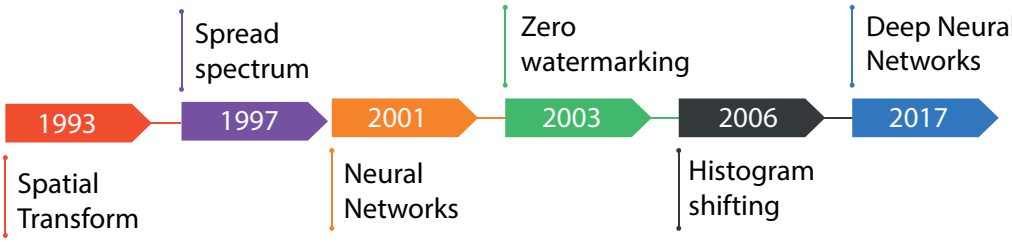


Figure 1. Watermarking operations chronology.

Digital image watermarking is a technique in which a secret mark is concealed in the host multimedia to create the watermarked image, which is then communicated through

a physical transmission channel. At this point, digital content may easily be retrieved, reproduced, and distributed illegally. Later on, the watermark is detected or extracted from the recovered image, which aims to confirming the ownership and guarantee tamper resistance. Watermarking schemes are divided into blind and non-blind based on the extraction requirements [3]. In blind watermarking, the extraction is achieved from the received container, whereas in non-blind watermarking, the original container is required for extraction.

For the purpose of preserving the content security of digital media, several techniques have been regarded as effective and practical techniques of copyright protection, such as cryptography and steganography [4]. The main goal of cryptography is to scramble the message, making it unintelligible and not subject to encryption or interpretation, in order to achieve the major security components. However, once the secret data have been cracked in some way, it will be entirely controlled by the attacker, and the original user will not be able to track it, especially if it has been cloned. Steganography, on the other hand, is considered as a one-to-one communication method based on invisibly hiding the secret data from unauthorized users by putting them into the cover media, so that only the sender and the intended receiver can identify the delivered message,; therefore, it can be used as an alternative tool for privacy and security. Yet, this technology is still not robust against attacks where a malicious manipulation or distribution of data makes the message easily lost and hard to recover. Subsequently, the emergence of digital watermarking technology perfectly solves the shortcomings of the last two defined techniques [5]. It is a technique similar to steganography; however, it overcomes the weaknesses in both cryptography and steganography. The concealed watermark is imperceptible and difficult to remove by unauthorized persons. The embedded watermark can be used to verify ownership if the container is targeted or tampered with [6]. This advantage makes image watermarking appropriate for a variety of applications [7].

In recent times, popular applications of watermarking [8–13] include 5G communication, Internet of Things (IoT), cyber systems, IP protection, relational databases, 3D images, fingerprinting, e-governance, digital forensics, military, medical, file archiving, hardware protection, secure social data in smart city, secure cloud storage data, e-voting, and remote education. A set of watermarking applications are shown in Figure 2. With the growth and popularity of the Internet, as well as the introduction of numerous live broadcast platforms, all types of pirated videos are flooding online. This seriously violates the rights and interests of owners of video copyrights and prevents the video industry from growing in a healthy way. As a result, robust video watermarking algorithms for copyright protection have developed to meet the actual needs [14]. Cyber–physical systems, the Internet of Things (IoT), the on-demand availability of computer system resources, and cognitive computing are four key elements of Industry 4.0. The effectiveness of this industrial revolution depends on how successfully these parts can cooperate and communicate with one another. This communication is accomplished by sharing a wide range of data acquired from a network of sensors, attracting the unwanted attention of hackers. As a result, one of the key concerns is data protection, and improved watermarking techniques are needed [9]. The ever-increasing volume of medical digital images, as well as the necessity to distribute them among specialists and hospitals for better and more accurate diagnoses, necessitate the protection of patients' privacy [10]. Important information is concealed within a cover medical image, and it should not be detected, retrieved, or modified by an unauthorized user. The challenge is high, and medical-image watermarking must be performed with extreme caution. The watermarking technique must not degrade image quality, and personal patient information included within the image must be retrievable without error following image decompression [15].

To provide services to residents, smart cities make use of information and communication technologies (5G, AI, cloud, and edge computing). A huge amount of information is collected, transmitted, and analyzed from each point within a smart city. This increases the vulnerabilities and risks of the acquired data [16]. Data validation is frequently required

in order to identify the location and technical details of how and when the data are generated. Improving their security is critical because malicious exploitation and piracy have become recurrent phenomena, imposing the need for digital content protection. Smart City Security is a collaborative project with many partners who are critical components of a complete, vital, and interconnected IoT security ecosystem. The present challenge is to identify smarter protection approaches and define a uniform framework for data-use policies. Digital watermarking techniques are highly advocated [16,17].

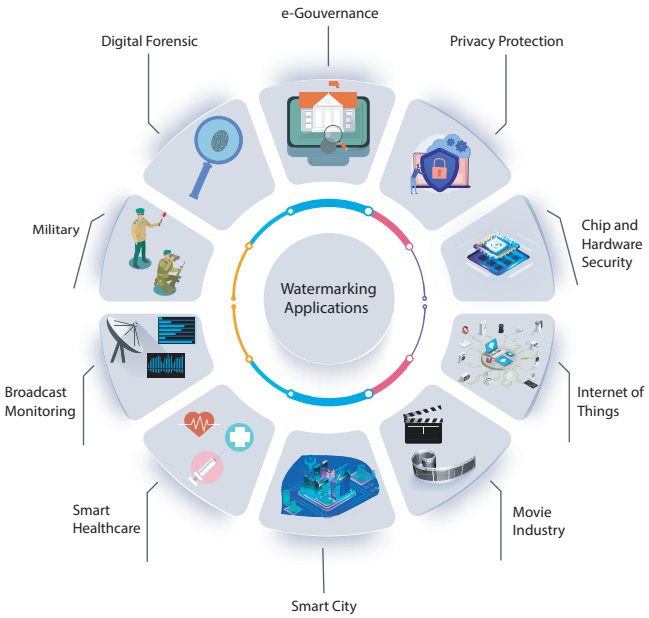


Figure 2. Recent applications of watermarking.

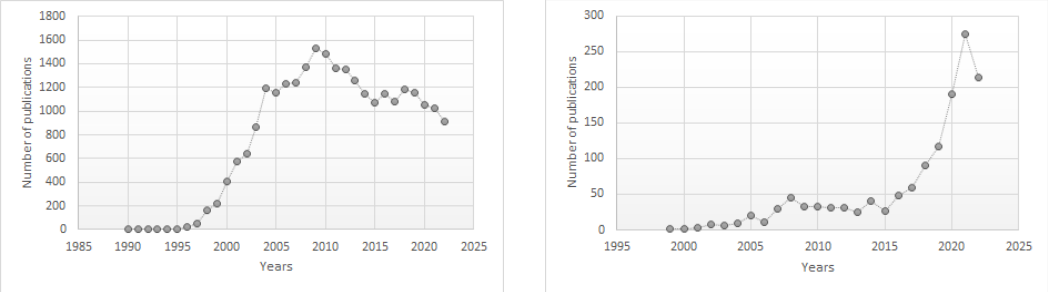
Watermarking techniques have been implemented by various algorithms employing spatial and frequency domains, as well as deep learning [18], with each having distinct benefits and limitations. Regardless of the field or the technology used, there are three common and mutually restricted properties in a watermarking scheme: invisibility, robustness, and watermark capacity or information quantity to insert. These three constraints are contradictory. For example, increasing watermarking strength to make the watermark more robust will have the opposite effect of making the watermark more visible. In the same way, increasing the capacity will result in more noticeable watermark that is also less robust. Therefore, we have to respect a compromise between these three criteria to build an acceptable watermarking method [19]. Yet, there are other significant properties related to watermarking algorithms, such as security and computational complexity. Overall, achieving a good relationship along with other watermarking properties is difficult.

Deep learning (DL) has attracted considerable attention in the field of digital image processing, and it has achieved great success in improving prediction performance by analyzing and learning from massive data and computational resources [20]. Furthermore, problems considered to be unsolvable are now solved with extraordinary precision. Likewise, digital image watermarking is a prime example of deep learning's significant contributions [21], giving rise to a large number of deep learning-based watermarking schemes with substantially better performance compared with traditional methods. However, deep learning will not solve all digital watermarking problems. There are some problems where traditional techniques with global features are a better solution. The advent of DL can pave the way for complex problems to be solved, or it can be combined with traditional techniques to fill in gaps and achieve better results [22,23]. To date, we have covered several concepts that are generally interconnected. However, it remains unclear how deep learning has been integrated into the digital image watermarking, what contri-

bution it has made to the field, and whether deep learning can really replace traditional techniques-based watermarking.

Several surveys on watermarking methods were proposed in the literature [7,21,24,25]. To the best of our knowledge, no work has been performed to investigate the transition of image watermarking from traditional techniques to ones based on deep learning. In this paper, we intend to explore current trends in digital image watermarking techniques. We will discuss two aspects of image watermarking. The first includes traditional techniques, and the second depends on deep learning architectures. We believe this classification will be useful in understanding the transition from traditional methods to deep learning-based methods. Alternatively, by comparing image watermarking schemes collected in the upcoming summary tables, researchers could gain significant insights into the image watermarking field with its new (modern) and old (traditional) technologies.

Figure 3, shows the number of documents (conference papers, articles, reviews, book chapters, etc.) by year according to Scopus data, using watermarking as a keyword in Figure 3a and watermarking-based learning in Figure 3b. As can be seen, after a period of stagnation between 1990 and 1995, the number of watermarking publications increased and reached a peak in 2010. On the contrary, watermarking-based learning publications began at the end of the 1990s, fluctuated between 2000 and 2015, and have since increased. In this review paper, the selection of references was performed as follows: Before 2015, the selection was based on a compromise between the paper's popularity (citations, journals), and its easy readability using classical keywords (watermarking, steganography, data hiding, etc.). After 2015, the selection was based on the popularity and variability of the proposed ideas and novelties based on additional keywords (machine/deep learning and watermarking).



(**a**) Watermarking documents. (**b**) Watermarking-based learning documents.

Figure 3. Documents per year based on Scopus website.

The following are the important contributions of this work: (1) To begin, we introduce a comprehensive taxonomy of digital image watermarking schemes based on several concepts and criteria. (2) Then, we present a recent state-of-the-art watermarking approach based on working domains, and a summary of the research results is described in tabular format. (3) Afterwards, we review the deep learning-based image watermarking methods and summarize and compare the contributions of the reviewed approaches in different technical perspectives. (4) Subsequently, we classify most popular reviews from the earlier sections according to how resilient they are to different forms of attacks. This classification aims at providing a clear vision of the strengths and weaknesses of deep learning technologies compared with traditional ones. (5) Finally, we state the main issues and challenges, as well as potential future directions.

This paper is organized as follows: Section 2 describes background concepts about digital image watermarking, including watermarking phases and different type of attacks. Section 3 presents digital image watermarking requirements with objectives measurements used to evaluate watermarking schemes. Section 4 reviews related algorithms on image watermarking. Furthermore, we discuss the current state-of-the-art image watermarking-based traditional techniques. Our reviewed techniques are also summarized in this section. Section 5 presents a comprehensive review and summary of modern image watermarking

methods based on various deep learning networks. Section 6 presents some issues and challenges, as well as the potential solutions to existing problems. We end with a conclusion and significant future directions in Section 7.

2. Image Watermarking Background

Digital watermarking is the process of embedding an invisible mark into digital data, such as an image, audio [26], a video [27], or text [28], etc. This mark is a random sequence of bits, a binary logo, or a message, depending on the application. Information is embedded in digital data by making invisible changes to the content of the data. Watermark detection must be robust even if the watermarked documents are attacked [29]. A general image watermarking scheme is shown in Figure 4.

Figure 4. General scheme of digital watermarking.

There are three major stages to digital watermarking. The insertion phase, the transmission phase (in which attack occurs), and the detection phase (or extraction). During the insertion phase, the mark is inserted into a digital document. After this phase, the original digital document is slightly modified—the modified document is referred to as a watermarked document. In the extraction phase, the mark is extracted from the watermarked document. The extracted mark is then compared with the original mark: if the two marks are identical then the document is authenticated, otherwise, the document is deemed falsified. During the transmission of the watermarked document over a public network, for example, it may be subject to certain attacks, which can alter the document [30]. If this alteration is significant enough, it will result in poor decisions.

2.1. Embedding Phase

In this phase, the original image I_0 is combined with the message m, which is a collection of bits encoding the data to be added. This operation's goal is to create a watermarked image I_w. The watermarked images are perceptually identical to the original images. The message m is encoded using a secret key K to generate a mark W. To maintain the imperceptibility requirement between I_w and I_0 [31], the mark is then modulated and/or scaled. There are two ways to insert a mark W into an image I_0—additive or substitutional.

2.1.1. Additive Method

The additive approach consists of adding the watermark to the image's components using one of the mathematical equations listed below [32]:

$$\begin{cases} I_w = I_0 + \alpha \times W \\ I_w = I_0(1 + \alpha \times W) \\ I_w = I_0 \times e^{\alpha \times W} \end{cases} \quad (1)$$

where α controls the watermark insertion strength to maintain the balance between the imperceptibility and robustness requirements. Watermarks can be added directly to the pixels of the original image or to the frequency components after image transformation, such as discrete wavelet transforms [33], discrete Fourier transforms [34,35], discrete cosine transforms [36], etc. Figure 5 shows the principle of additive insertion.

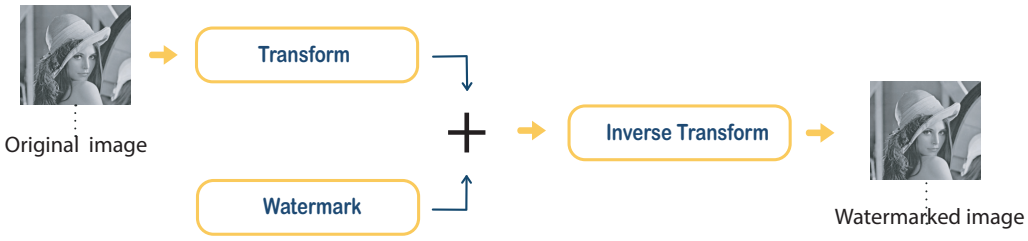


Figure 5. Embedding by additive method.

2.1.2. Substitutive Method

In this case, the mark to be inserted is substituted for components of the original image. There are several substitution methods, such as the substitution of the least significant bits (LSB) [37], which consists of replacing the least significant bits of the pixels of an image by the bits of the mark. Another technique widely used in image watermarking is quantization by index modulation (QIM), which is proposed in [38]. The basic idea behind this technique is to quantize the image using the quantizer that corresponds to the mark to be inserted [39]. Thus, each element of the mark is associated with a different quantifier.

2.2. Detection Phase

The image watermarking system aims to permanently embed data in the original image and then attempt to accurately identify or extract it. The watermarked image I_w can be subjected to different attacks, resulting in a distorted watermarked image I_w^*. Watermark detection/extraction from the distorted watermarked image I_w^* must be conducted during the detection phase. According to different image watermarking schemes, there are two modes of watermark detection/extraction. The first mode requires the original image and the second does not [32]. The choice of the mode will depend on the application envisaged and the protocols employed.

2.2.1. Blind Watermarking

The watermarking technique is qualified to be blind if the hidden information is detected or extracted directly from the watermarked image without requiring the original image [40]. Blind watermarking is widely preferred in most applications since the original image is usually not available at the stage of detection/extraction.

2.2.2. Non-Blind Watermarking

The watermarking technique that requires the original image to extract the mark is called non-blind watermarking [41]. It is more robust because the watermark can be easily detected or extracted using both the original image and the watermarked image; however, non-blind watermarking is less popular because it requires the original image.

Figure 6 shows the watermark detection/extraction of additive watermarking. The watermark characteristic components are extracted from the watermarked image. The key K is used to generate the reference mark. The mark is then decoded, either by estimation or by correlation of the components, to produce a measure indicating the probability of the presence of the mark in I_w^*.

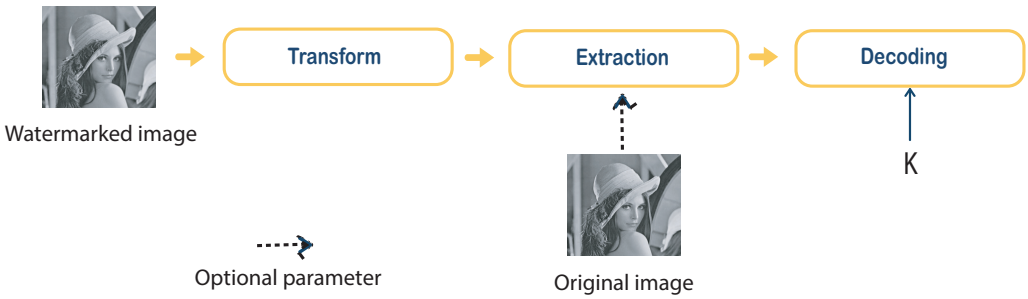


Figure 6. Detection by additive method.

2.3. Transmission Phase

The watermarked images are vulnerable to several attacks during the transmission phase, which is the intermediate phase between the embedding and detection/extraction phases [32]. Attacks in digital watermarking are defined as the set of operations or modifications that can make the watermark undetectable. There are two major categories of attacks:

2.3.1. Unintentional Attacks

Any processing or manipulation performed by a non-malicious user is classified as an unintentional attack. The goal of these treatments is to damage the watermark by modifying or hiding some characteristics of the image but not to remove the mark. We will content ourselves to provide the most frequently encountered examples.

Compression: Compression attacks are designed to reduce the amount of data encoded in the host document. This process consists of removing the perceptibly less important components while preserving the important components. JPEG and JPEG2000 are considered the most popular image compression algorithms currently in use [42].

Filtering: Generally, filtering is used for noise reduction. This type of filtering basically has the effect of attenuating the high-frequency components of the image and therefore degrading the brand components inserted in these frequencies.

Volumetric transformations: The basic idea behind these transformations is to adjust the brightness value of each pixel in an image using a linear or non-linear function to improve the visual aspect of the image. These transformations include histogram spreading and equalization, Gamma transformation, and so on.

Noise: The addition of noise increases the uncertainty or ambiguity of the average information embedded in the image, which can have a masking effect on the watermark and therefore disturb its detection/extraction.

Geometric transformations: This kind of transformation does not remove the mark, rather, it makes it undetectable while still remaining in the image. In the majority of watermarking algorithms, the watermark detector requires knowledge of the exact position of the mark in the image. These transformations cause a desynchronization between the mark inserted in the image and the detector [43]. There are several geometric transformations, for example, affine geometrical transformations (translation, rotation, and change of scale), and cropping, which consists of removing a part of the image and consequently a part of the mark. There are also local geometric transformations, such as the StirMark attack. This attack consists of a succession of random geometric distortions applied locally to several places in the image [44].

2.3.2. Intentional Attacks

These are specific manipulations meant to intentionally destroy or fake the watermark in order to prevent its extraction/detection. All of the attacks presented above can be used for this purpose. We also mention:

Cryptographic attack: The main goal of this attack is to crack the security codes and algorithms used in watermarking methods in order to remove the inserted watermark

information. In this case, the attacker tries to identify the embedded secret key using exhaustive search techniques. This type of attack possesses high computational costs and it is limited, and therefore it is employed less frequently.

Protocol attack: The idea behind this attack is to create document ownership ambiguity. Some protocol attacks are based on invertible operations, in which an attacker subtracts their watermark from the watermarked document and claims its ownership. Another type of protocol attack is the copy attack [45]. In this case, the goal is to remove the watermark from one container and place it in another target document. In all cases, the attacker can claim for the ownership of both the original and watermarked image.

The second type of attacks is more dangerous. It primarily relates to system security. However, this type of attack is difficult to perceive and thus is omitted in most of the works reviewed. The attack of the first type is considered successful if the watermark is damaged. Accordingly, several counter-techniques have been developed by embedding in the most prominent features of the container. This goes against the concept of improving watermark imperceptibility. Thus, striking a balance between imperceptibility and robustness is a major challenge in creating a watermarking outline.

3. Performance Requirements in Digital Watermarking

To have a good watermarking scheme, the watermarking design algorithm must meet some requirements and characteristics [46]. These requirements evaluate the performance of the watermarking technology. The significance of each property depends on the intended watermarking application. In the following, we present the main requirements for a digital watermarking scheme with the corresponding evaluation metrics used in each case.

3.1. Robustness

This measures the capability of the hidden watermark to resist any signal-processing manipulation. These modifications define all attacks, whether intentional or non-intentional. The first type of attack is aimed at damaging the invisible watermark, while the second type of attack is not explicitly aimed at modifying the watermark but rather at removing it. The level of robustness varies depending on the watermarking approach used, and certain techniques are effective against some attacks but fail against others. Based on robustness criteria, three types of digital watermarking are identified:

Robust watermarking: In this case, the watermark must be resistant to different attacks on the digital document, and the detection of the watermark must be effective even under those manipulations [47]. This watermarking technique is suitable for a variety of applications, including copyright protection, fingerprinting, broadcast monitoring, and copy control [48].

Fragile watermarking: In fragile watermarking, the mark is extremely sensitive to the modifications of the watermarked document [49,50]. This technique is used to prove the authenticity and integrity of multimedia data. A fragile watermarking technique is designed to detect (with a high probability) any kind of manipulation of the watermarked document, including both incidental and intentional attacks. A comparison of the extracted watermark and the original watermark is performed to identify if the document is modified or not.

Semi-fragile watermarking: This type of watermarking combines the characteristics of robust and fragile watermarking into an intermediate situation in which the watermark is robust to a defined set of attacks but fragile to others [51,52].

Generally, two indices are mostly adopted to measure the robustness of watermarking methods:

Normalized Correlation (NC): NC computes the similarity and difference between the original watermark and the extracted watermark. The NC value of a good watermarking algorithm should be $\geq$0.7. Ideal algorithms provide an NC equal to 1. NC is defined as follows:

$$NC = \frac{\sum_{i=1}^{X} \sum_{j=1}^{Y} (W_{org_{ij}} \times W_{ext_{ij}})}{\sum_{i=1}^{X} \sum_{j=1}^{Y} W_{org_{ij}^2}} \quad (2)$$

where $W_{org_{ij}}$ and $W_{ext_{ij}}$ are the original watermark and extracted watermark, respectively, of size $X \times Y$.

Bit Error Rate (BER): BER is also one of the measurements that evaluate the robustness of a watermarking technique. BER computes the ratio value of incorrectly extracted watermark bits from the inserted watermark bits. The closer the BER value is to 0, the greater the similarity between the extracted watermark and the original one, indicating more robustness.

$$BER = \frac{NI}{NT} \tag{3}$$

NI is the number of incorrectly extracted bits, while NT is the total number of bits transmitted.

3.2. Imperceptibility

The concept of imperceptibility is related to the visual system or auditory perception of watermark insertion in the host document. Therefore, in the case of images, the watermark should be invisible to a human observer in such a way that the inserted watermark does not affect the quality of the document [53]. The evaluation of the imperceptibility after watermark embedding is an important criterion for watermarking algorithms validation. For images, such an evaluation requires human visual system (HVS) analysis. Several objective measures have been proposed in the literature to measure the visual quality of a watermarked document.

Peak signal-to-noise ratio (PSNR): PSNR is the most commonly used indicator for providing quantitative scores across the watermarked images. PSNR evaluates the invisibility requirements by comparing the similarity of the original image file to the watermarked one. The PSNR is defined as:

$$PSNR = \log_{10}\left(\frac{d^2}{MSE}\right) \tag{4}$$

where d is the maximum pixel value of the image and the maximum possible value for a pixel. For example, this is $d = 255$ if the image pixels are coded with 8 bits. MSE represents the mean square error, which is defined for two images using the following equation:

$$MSE = \frac{1}{M \times N}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}(I_0(x,y) - I_w(x,y))^2 \tag{5}$$

where I_0 and I_w are the original and watermarked image of dimension $M \times N$, respectively. The higher PSNR value can be obtained with a low MSE value, indicating low degradation. In general, if the PSNR is greater than 40 dB, the watermark is considered invisible, while a value below 30 dB indicates significant distortions [54].

Structural similarity index model (SSIM): The SSIM evaluates the structural similarity existing between two images. Unlike PSNR, which is based on the pixel-to-pixel difference between two images. The SSIM measure consists of combining three parameters: brightness, contrast, and structure comparison. It is computed on several windows of an image. The measure between two windows x and y is given by the following:

$$SSIM(x,y) = [l(x,y)]^{\alpha}.[c(x,y)]^{\beta}.[s(x,y)]^{\gamma} \tag{6}$$

where $l(x,y)$, $c(x,y)$, and $s(x,y)$ represent the luminance function, contrast function, and structure comparison function, respectively. x and y are the pairs of local square windows of the original and the watermarked image, respectively, and the importance of each measurement is defined by the parameters α, β, and γ. Finally, the similarity between

the reference image and the watermarked image is measured by the average of the SSIM calculated for each window and denoted by MSSIM:

$$MSSIM(I_0, I_w) = \frac{1}{m}\sum_{i=0}^{m} SSIM(x_i, y_i) \tag{7}$$

where I_0 and I_w are the original and the watermarked, respectively, and m represents the total number of windows. The range of MSSIM is from 0 to 1; therefore, the MSSIM value must be close to 1 to show high similarity between the two images.

3.3. Capacity

The capacity of a watermarking method represents the highest amount of information that can be hidden in the cover document while maintaining the imperceptibility property. This quantity of information depends essentially on the type of intended application [55]. The information is commonly represented as bits per pixel (bpp). The capacity value can be defined as follows:

$$Maximum\,embedding\,capacity = \frac{Number\,of\,pixels\,with\,peak\,point}{image\,length \times image\,width} \tag{8}$$

For example, high watermark robustness is achieved at the expense of robustness, imperceptibility, or both. Similarly, the more information that is inserted, the more distortion appears on the host image. As a result, a good trade-off has to be maintained between these types of properties. Figure 7 illustrates the trade-off among these three requirements.

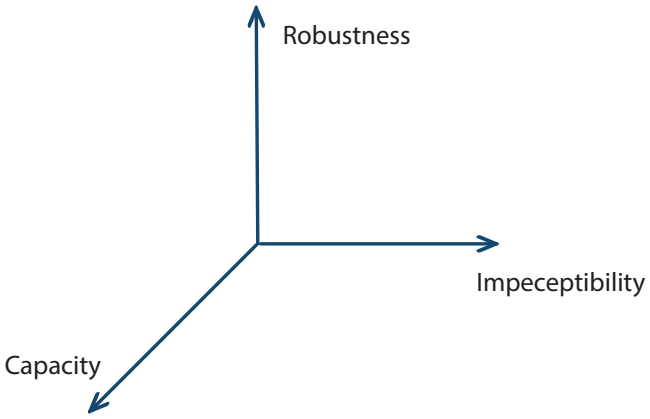


Figure 7. Constraints of digital watermarking.

There are other properties related to watermarking algorithms that must be taken into account, such as complexity, security, and false positive rate.

3.4. Complexity

The complexity property is a critical element for real-time watermarking applications. It describes the computational cost of embedding the watermark into a host image, as well as detecting/extracting the watermark from the watermarked image. In general, the computation time must be as short as possible and less than a certain value [56]. For example, in access control, the complexity of the insertion is less important than the complexity of the detection. The computation time of the detection must be of the order of one to two seconds. The complexity can be defined by computing the embedding and extraction times.

3.5. Security

The security issue refers to the watermark's ability to resist attacks that intend to thwart the watermark's purpose. The watermarking scheme is considered to be secure if it

is difficult to remove or alter the watermark without causing damage to the cover image or without knowledge of the watermarking embedding and detector [57]. Encryption methods, such as DCT- and chaos-based have been used to ensure watermark security, where the encryption key identifies the degree of security [58,59].

Number of Changing Pixel Rate (NPCR) and **Unified Averaged Changed Intensity (UACI)**: Those two measures are used to evaluate the efficiency of image watermarking against potential attacks. They are usually used to analyze the resistance of the watermarked images to pixel-level changes. NPCR and UACI scores should always be close to 1 and 0.33, respectively, to achieve good security. Higher values mean resistance will be better. If W and $\overline{W}$ denote the original and the extracted watermark, respectively, NPCR and UACI are defined as:

$$NPCR = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} D(i,j)}{M \times N} \times 100 \quad (\%) \tag{9}$$

where:

$$D(i,j) = \begin{cases} 1 & if\ W(i,j) = \overline{W(i,j)} \\ 0 & Otherwise \end{cases} \tag{10}$$

$$UACI = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} |W(i,j) - \overline{W(i,j)}|}{M \times N \times 255} \times 100 \quad (\%) \tag{11}$$

3.6. False Positive Rate (FPR)

FPR represents the probability that the detector indicates the presence of the watermark when, in fact, no watermark is inserted or a wrong watermark, which has never been inserted into the host image, is extracted [60,61]. The main reason for the false positive error is the dependence of the extracted watermark on information provided by the user. This measure should be low for any watermarking system, and it can be defined as follows:

$$FPR = \frac{FP}{TN + FP} \tag{12}$$

where FP and TN are the total numbers of false positives and the number of the true negatives test results, respectively.

4. Conventional Image Watermarking Schemes

According to the domain of insertion, the watermarking techniques proposed in the literature can be grouped into two classes: those that operate in the spatial domain and those in the transformed domain [7]. A watermark can be embedded in the spatial domain by involving a matrix of pixels in a digital image [62], and it can also be embedded in the frequency domain by modifying the frequency spectrum of an image generated through some type of frequency transformation [63]. However, there are techniques that combine several domains to achieve better watermarking performance [64]. The robustness and capacity of the approach are directly impacted by the domain selection. This choice depends on the targeted application, since each domain has its own characteristics. Further, in this section, we review various works in the most widespread domains of digital image watermarking.

4.1. Spatial Domain

It is easier to work with the spatial domain than with the frequency domain because no transformations are required. At the same time, the frequency domain is more robust since it allows obtaining an image representation that is invariant to slight modifications, resulting in better robustness of a watermark [65] but worse imperceptibility. In techniques based on the spatial domain, such as as the insertion domain, the mark is inserted by a direct modification of the pixels. These pixels are usually selected using a secret key or are based on a psycho-visual model. In the most commonly used designs, the values of the color channels, luminance signals, or brightness signals of the digital image are used in this

situation [66], based on the logo, signature, or random bits defined by the owner. Many watermarking techniques are proposed in this regard, such as least significant bit (LSB) modification algorithms, intermediate significant bits (ISB) and patchwork algorithms, spread spectrums, and correlation-based algorithms. Because of their simplicity, the earliest digital watermarking algorithms were developed to work in this domain. Tanaka et al. [67] proposed the first spatial watermarking method. It consists of inserting each element of the mark in the least significant bit of each pixel. Another method called the "patchwork" algorithm has been proposed by [68]. This algorithm also operates directly in the spatial domain. The watermark is concealed with a specific Gaussian distribution in the luminance values of the original image. To retrieve the watermark, correlation-based algorithms are mostly used in the spatial domain.

In the following, we review several recent watermarking approaches based on the spatial domain that have attracted the attention of the research community in the last few decades because of their optimal balance among robustness, imperceptibility, and capacity, all of which are required for any watermarking technique.

First, Thongkor et al. [69] presented a digital watermarking method for camera-captured images. In this scheme, they embed a binary image with the same size as the host image, with each pixel in the host image used to carry a watermark bit. Several types of pixel values distortions and geometric distortions were investigated after the watermarked image was printed and captured by a digital camera. The method shows robustness against various types of distortions from the printing and camera-capturing processes. However, with the distortions introduced, reliable extraction was not achievable on certain levels, essentially because of camera position, camera settings, and lighting conditions.

Desynchronization attacks produce a delay between the insertion and detection algorithms. As a result, the watermark often has difficulty surviving this type of attack, especially for color images. Pan-Pan et al. [70] proposed a new robust color-image watermarking method based on a significant bit-plane histogram. Here, the watermark is embedded in the affine invariant local feature regions of the RGB plane. The scheme is invisible and robust against common signal-processing and desynchronization attacks. However, the method has a lower watermark capacity and a higher computation time, making it unsuitable for real-time applications.

Belferdi et al. [71] designed an effective fragile watermarking scheme for RGB image tamper detection and restoration. In this technique, three copies of the watermark are concealed over three components (red, green, and blue channels) of the host image, improving detection accuracy and restoration capabilities. The Bayer pattern is used to decrease the capacity of insertion by converting the host color image into a gray-level watermark. Moreover, the Torus automorphism permutation is adopted in order to scramble the gray-level watermark. However, security analysis in this work is inadequate, and the method must to be used in real application scenarios where security issues become a serious problem.

In contrast to conventional spatial domain watermarking, the embedding procedure is carried out in [72] without significantly degrading image quality or changing perceived color. To ensure high robustness against attacks and high image quality, the watermark is spread to wider parts of the image. The current scheme conceals a logo image in the color image based on two masks, embedding mask $M1$ and compensation mask $M2$, where the last mask ensures that the original color distributions are least affected by the embedded bits. In addition, the use of masks ensures that the modified pixels are not noticed compared with the neighboring pixels. Results reveal the high quality of the watermarked image and its high resilience to attacks, although only the blue component is used for embedding. Having no distinction between the host image and the watermark image is the worst-case situation. Furthermore, the LSB technique is easily altered, which affects data integrity and safety.

In [57] a key-based authentication scheme is described. The SHA-1 (secure hashing algorithm) algorithm is used in this case to compute the hash value for each image block, and the hash-key code is then generated and embedded in the block itself using LSB

substitution. The watermarked image shows high PSNR and SSIM values during testing. For tamper detection, the extraction process is similar to the integration process, and the same blocks and keys are generated. At this point, the extracted bits are the hash-key code that was hidden during the embedding process. If the generated code is similar to the embedded one, it is assumed that the block has not been tampered with. The watermarking technique was able to detect different tampering shapes and sizes, and it showed resistance to various image tampering attacks, such as copy–paste, copy–move, and constant average attacks. Nevertheless, the scheme does not incorporate robustness and self-recovery features.

In order to handle the computational cost, Gull et al. [73] elaborated a fragile watermarking technique in the spatial domain for the tamper detection and localization of medical/general images. In this approach, the cover image has been subdivided into two blocks called upper pixel block (UHB) and lower pixel block. From the watermark data, two-bit streams were produced; one is embedded in the LHB and facilitates tamper detection, while the other is embedded in the UHB and aids in localization.

The proposed scheme had low computational complexity, which makes it a better candidate for medical image authentication in real-time applications. In this case, the watermarking system is based on dividing the medical image into two regions, which are known as regions of interest. This type of watermarking is fragile, which means that it cannot withstand unintended attacks; moreover, irreversibility is the fundamental shortcoming of this work.

More recently, Rinki et al. [74] introduced a novel digital watermarking system using a matrix multiplication-based LSB mechanism. The algorithm selects the non-consecutive matrix of the pixels to substitute the last three bits of each RGB component of the host image with the gray-scale watermark image. In order to select an appropriate block for the embedding process, the filtering process is performed on each matrix block. The opposite of the embedding procedure is used to extract the watermark. The watermark robustness is tested against several types of attacks, such as salt and pepper noise and geometric transformations. However, the proposed method is non-blind, which requires the original image at the destination side for the extraction process.

Similarly, the methods used in each technique vary depending on the target application. In order to achieve the trade-off between robustness and short running time, the R matrix of QR decomposition is adopted in [75] to protect the copyright of color images in the spatial domain without true QR decomposition. Furthermore, Mustaqim et al. [76] proposed a robust watermarking scheme in the spatial domain to minimize the security vulnerabilities in the watermarked image. The fundamental idea behind this work is to embed a compressed color watermark into the luminance channel of the host image using a quantization process with pseudo-random steps. Both methods show good results in terms of robustness and time complexity. In general, watermarking methods based on the spatial domain remain sensitive to several types of attacks, especially synchronization attacks.

The primary advantages of spatial domain-based methods are their low computational cost, large watermark capacity, and improved efficiency, which favor their use in real-time watermarking applications. However, spatial domain techniques perform well only when the image is not subjected to any type of distortions, making them less robust to several attacks, such as noise and compression.

Table 1 summarizes an overview of the spatial domain-based state-of-the-art schemes.

Table 1. Summary of spatial domain-based image watermarking techniques.

Reference	Approaches	Objectives	Cover/ Watermark Size	Evaluation Metric	Advantages and Weaknesses
[70]	Invariant significant bitplane probability density	Robustness against desynchronization attacks	512 × 512 pseudo-random sequence of 85 bits	PSNR, MSE, SSIM	+ Resistance to common image-processing attacks and desynchronization attacks - High computational complexity - Less robust against JPEG compression and scaling
[73]	LSB, ISB	Tamper detection and localization of medical/general images	256 × 256 64 × 64	PSNR, SSIM, BER	+ Low complexity + Highly visual image - Fragile - Lack of irreversibility
[74]	LSB substitution mechanism	Enhance the conventional LSB technique of digital image watermarking	700 × 700 100 × 100	PSNR, NCC	+ Low complexity + High-quality image - Non-blind - Less robust against median filtering and speckle noise

4.2. Transform Domain

Usually, spatial watermarking techniques can be easily manipulated; therefore, these techniques are too fragile and less robust to different types of attacks compared with transform domain algorithms. These shortcomings have led to the search for transform domain watermarking techniques that successfully hide data in the frequency coefficients of the transformed image, rather than the spatial domain [77]. Watermarking methods in the frequency domain convert an image from the spatial/temporal domain to another domain using a pre-defined transform. Then, the watermark is inserted into the selected coefficients of the transformed image [78]. Finally, the inverse transformation is performed to retrieve the watermarked image. Figure 8 shows the steps of image watermarking in the transform domain. Digital watermarking algorithms designed to work in a transformed domain are more robust and complex; however, they are extensively employed [79]. Regarding frequency transformations, including the discrete cosine transform (DCT), discrete Fourier transform (DFT), discrete wavelet transform (DWT), and singular value decomposition (SVD), among others, the energy of the signal is concentrated in the low-frequency components. The insertion of a mark in these frequencies provides good robustness; however, it introduces apparent distortions in the spatial/temporal domain [80]. On the other hand, insertion in the high-frequency components does not degrade the signal quality but makes the mark vulnerable to attacks, such as low-pass filtering and compression. Therefore, the mid-frequency band is generally the best choice for watermarking, since it fits the requirement of compromise between robustness and invisibility. Several studies on image watermarking based on the transform domain have been carried out, showing good robustness, imperceptibility, and security to different attacks, such as image-processing attacks and geometric transformations. This section reviews the most recent of these studies that primarily use frequency domain transforms, such as DCT, DFT, and DWT.

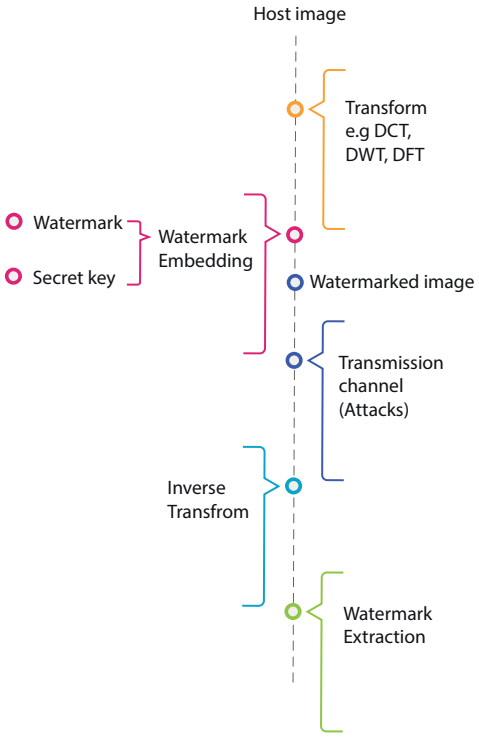


Figure 8. Watermark insertion and detection in the transform domain.

4.2.1. Discrete Cosine Transform (DCT)

DCT converts the image into equivalent frequency coefficients by modulating the frequency components. These coefficients are distributed into three regions: low, medium, and high frequency. After converting the image into the DCT domain, the left IP corner retains the majority of the image energy, which makes the DCT suitable for image compression. Moreover, it indicates a good balance between its near-optimal decorrelation and the lower complexity of the algorithm. DCT-based image watermarking methods have been the subject of many investigations.

In [81], blind removal-based dual image watermarking was proposed using DCT and quantization index modulation dither modulation (QIM-DM). First, the watermark pattern generated from the binary watermark image is embedded into the luminance component using two secret keys, resulting in the dual watermarked image. In the extraction stage, the same two keys are employed to extract the concealed watermark. The proposed work ensures a blind removal, good image quality, and robustness to several intentional and unintentional attacks, including JPEG compression. However, using incorrect keys during the removal process will greatly distort the watermarked image.

Likewise, Loan et al. [82] proposed robust and secure color-image watermarking in the DCT domain for grayscale and color images. The watermark is embedded by modifying two preselected DCT mid-frequency coefficients of adjacent blocks. The Arnold transform, in addition to chaotic encryption, are used to provide security in this case. Their experimental findings revealed robustness against image-processing operations, such as compression, cropping, and median filtering. Despite this, although the discussed scheme is secure and robust against several types of attacks, including image-processing operations, it will not be suitable for some applications due to its high complexity value.

In [83], a semi-blind DCT watermarking method was proposed based on differential evolution and a kernel extreme learning machine (DE KELM). Entropy is used to select non-overlapping blocks. Then, DCT is applied to the selected block. The low-frequency coefficients are selected in a zig-zag manner, considering the quality of the watermarked image. In addition, KELM used the non-linear regression model to predict the embedding coefficient for each block, while DE is used to control the watermark strength. This

watermarking scheme provides robustness to a variety of attacks, such as JPEG compression, histogram equalization, and Weiner filtering. However, it is fragile to rotational attacks. Additionally, the security and payload can be improved, and the KELM parameters can be adjusted for better imperceptibility.

On the other hand, Moosazadeh et al. [84] designed a watermarking scheme based on DCT for JPEG-YCbCr images. DCT is performed on the "Y" component, and the complexity of each generated block is computed. Then, the most complex blocks are selected for watermark embedding. In addition, both DCT coefficients and YCbCr color space are stable against most changes in the host image, which increases the scheme balance. Here, Teaching–Learning-Based Optimization (TLBO) will generate the embedding parameters and the most suitable position for concealing the watermark, thus ensuring better imperceptibility of the watermarked image and high robustness against various distortions. Nevertheless, the proposed scheme is vulnerable against median filter attacks and JPEG compression. Moreover, the computational cost can be reduced.

In order to achieve a good trade-off between robustness and invisibility, Wang et al. [85] recommended a color-image watermarking scheme using discrete cosine transform (DCT) based on just-noticeable distortion (JND) with Arnold transform. Here, the watermark is scrambled using the Arnold transform, then embedded into the DCT coefficient of the "Y" channel of an image in the YCbCr color representation. Furthermore, the proposed scheme includes a novel color complexity weight from the "Cb" channel for high robustness. After evaluating the watermarking system under different types of attacks, such as JPEG compression, Gaussian noise, and geometric attacks, the superiority in terms of invisibility and robustness is evident. However, it should be highlighted that the time complexity analysis deserves further investigation, especially since JND has been used in the scheme.

For the same purpose, Zhang et al. [86] recently designed a perceptual image watermarking scheme based on a tri-directional correlation of DCT coefficients. After generating these coefficients, the difference between two DCT blocks is determined based on three directions in the neighborhood. Specifically, the watermark embedding is performed using dependencies between the middle-frequency coefficients of adjacent blocks based on the JND model. Although, the results demonstrated high robustness to common image-processing attacks, the robustness tests were insufficient.

Using the DCT transform in image watermarking techniques has various advantages, some of which have already been mentioned. However, geometric transformations can significantly alter the values of the DCT parameters, hence the interest in using a field with invariant properties against geometric transformations, such as rotation and translation. The different approaches dealing with desynchronization attacks reported in the literature are shown below.

4.2.2. Discrete Fourier Transform (DFT)

In the Fourier domain, the image is decomposed into a sum of elementary signals in the form of complex periodic Fourier coefficients and represented as amplitude and phase. DFT properties allow the watermarking scheme to handle most geometrical distortions. Therefore, several DFT-based image watermarking approaches have been proposed.

Firstly, Liao et al. [87] came up with a frequency watermarking technique based on DFT and compressive sensing (CS) to increase the quality of the recovered image. This method is based on separable data hiding in encrypted images. In the first step, the selected coefficient matrices are generated from the DFT cover image; then, they are divided into the most significant information used for encryption, while other information is used for embedding. The scheme demonstrated excellent recovery and provided a flexible payload. However, there is still a need to realize fully reversible data hiding. Moreover, CS algorithms are time consuming, which reduces the chances of adopting this global watermarking method in real-world applications.

In [88,89], the authors proposed a DFT-based robust watermarking method for printed and scanned ID images. The watermark is circularly embedded within the DFT coefficient's

magnitude. In this context, two mechanisms have been associated with DFT-based image watermarking. A Wiener filter is used to reduce image blurring, while color correction is used to attenuate color degradations. The whole scheme has good resistance to geometric attacks produced by the print–scan process, and it shows high efficiency and low computational cost. Further, Gourrame et al. [90] improved the previous watermarking scheme for the print–cam process by including a pre-processing stage to correct the perspective deformations using the Hough line transform. Accordingly, enhanced robustness against the strong geometric attacks produced by the print–cam process.

On the other hand, Prajwalasimha et al. [91] introduced a non-blind and Fourier-based image watermarking scheme. A bulk watermark is generated, then the DFT is applied to the cover image. Following that, a progressive division is applied to the bulk watermark in order to scale down the intensity of each pixel. To retrieve the watermark, the DFT watermarked image and DFT host image are subjected to a detection process based on successive multiplication processes. Better imperceptibility is observed in the results with less execution time. The algorithm provides the same data capacity as other algorithms studied in this work, and DWT can be used to increase data-embedding capacity. Likewise in [92], various variants of the Fourier transform were used (e.g., discrete Fourier transform, a fractional Fourier transform, as well as a quaternion discrete Fourier transform) to conceal the watermark in the medium-frequency band of the original image. According to the results, the system performs well against filtering attacks and produces high-quality watermarked images. However, the embedding scheme requires three coefficients to insert two bits, significantly reducing the capacity of the watermarking method.

Chen et al. [93] exploited DFT-based watermarking embedding to cope with perspective distortions produced by the screen–cam process. For this purpose, the watermark is concealed in the middle-frequency DFT coefficients, which are quite resilient to changes in media data. At the extraction phase, a synchronization mechanism is used to rectify the perspective transformations. Finally, the watermark is extracted using a message extraction algorithm based on the local maximum value from the DFT of the noise component. This method shows robustness to common attacks, as well as screen–cam attacks. Although perspective correction will never be perfect, the correction must enhance the effectiveness of the automatic perspective correction approach so that it may be employed in real-time copyright verification applications. Furthermore, the watermark capacity should be increased.

Further, Hsu zet al. [94] designed a blind color-image watermarking scheme using quaternion discrete Fourier transform (QDFT). The proposed method combines a number of algorithms, such as extreme pixel adjustment (EPA), mixed modulation (MM), multi-bit partly sign-altered mean modulation (MPSAM), and particle swarm optimization (PSO). In this approach, these techniques are collectively referred to as EMMQ. Due to the weakness of QDFT-based techniques against JPEG compression and contrast enhancement attacks, EMMQ has been employed to provide good robustness against these types of attacks. MPSAM is used to map multiple bits within a single non-overlapping image partition block. In contrast, the MM is used to embed the watermark so that the trade-off between robustness and image quality can be achieved, while PSO has been used as a swarm intelligence technique to optimize the EMMQ parameters, which generally optimizes the results in both robustness and imperceptibility compromises. Experiment results prove the resistance of the watermarking scheme to several image-processing attacks. However, the embedding and extracting operations required a high computational cost, and geometric attacks were not considered in this work.

In [95], the amounts of DFT bands and the frequency coefficients, as well as watermark strength, were optimized using particle swarm optimization. The visual quality metric called VIF was suggested to obtain better performance in terms of imperceptibility. In order to implement the algorithm in large-scale image datasets, 1000 color images with 24-bit/pixel resolution were used for testing imperceptibility and robustness. Additionally, PSO is employed in DFT watermarking to guarantee security by locating the specific secret key for each image in a given dataset. This approach also offers advantages in terms of

imperceptibility and robustness. Nevertheless, the compromise between imperceptibility and robustness needs to be improved; meanwhile, the use of VIF may result in significant dB-value loss for the images being examined.

More recently, Ref. [96] proposed a novel discrete Fourier watermarking method based on gray component replacement (GCR) masking for printed images. The watermark was concealed in the amplitude coefficients in the Fourier domain. The key innovation of this work is the use of a GCR mask to hide the artifacts produced by watermark embedding, which improved the visual quality of the watermarked images.

Due to the magnitude invariance characteristics, DFT-based watermarking techniques in general provide great resilience to a wide range of attacks, particularly geometric ones. As a result, the DFT domain is an important field of study. Higher watermark robustness, however, frequently produces perceptible visual artifacts. This makes it difficult to achieve a good balance between robustness and other watermarking requirements, such as imperceptibility and capacity.

4.2.3. Discrete Wavelet Transform (DWT)

This transformation decomposes an image into a low-frequency sub-band LL and three high-frequency sub-bands: LH, HH, and HL, which correspond, respectively, to the vertical, diagonal, and horizontal spatial orientations. The LL sub-bands can be further decomposed into k-levels. Therefore, one or more sub-bands of the same level or many levels can be used to embed the watermark depending on the targeted application and on the proceeded protocol. Diverse watermarking techniques developed on different levels for different purposes are presented.

In [97], the authors designed a robust and invisible image watermarking scheme in the wavelet domain using the quantization technique. Here, the grayscale watermark was first decomposed into binary images ordering from least significant bit (LSB) to most significant bit (MSB). A quantization method is then used to conceal the binary bits into the carrier's wavelet coefficients. For the extraction process, a classification threshold is defined using the Otsu algorithm to extract the watermark accurately. The suggested approach achieved high imperceptibility of the watermarked image and good robustness of the extracted watermark. However, it is still sensitive to geometric attacks, as well the high processing time of the algorithm makes it difficult to use in real-time applications.

The scheme in [98] introduced a robust non-blind watermarking algorithm for YCbCr color images based on channel coding. In the embedding stage, the original image was decomposed into four-level discrete wavelet transform (DWT). Then, the singular value matrices of the encoded watermark image were embedded into the Y, Cb, and Cr components of the image. For that purpose, just-noticeable distortion (JND) was used to conceal the mark in the sub-bands of the image, which provides an embedding strength. Experimentally, the scheme can achieve good transparency and robustness against various kinds of attacks. However, The coding rate of repetition code has a significant impact on the scheme's effectiveness. Furthermore, the non-blind approach is not the best option for some real-time applications.

In [99], the invariant integer wavelet (IIW)-based watermarking technique was presented with a new two-level encryption algorithm. The host image is transformed into the frequency domain using the redistributed invariant integer wavelet transform (RIIWT), and QR and Singular value decomposition (SVD) are used to find the singular value of the image matrix. Afterward, The watermark is concealed in the low and high frequencies of the IIW-host image. The present scheme provides rotation stability and lowers computational cost given the use of IIW and QR decomposition tracking by SVD, respectively. Moreover, it demonstrates significant resistance to geometric attacks and standard image processing, including different filtering and cropping attacks. At the receiving part, two watermarks are obtained by inverse RI-IWT on the low-frequency sub-band. Then, the most robust watermark with a higher NC value is taken as the final watermark, However, it is worth noting that it is not practical for all applications.

The wavelet-packet transform (WPT) is a type of wavelet-based transform that differs from the DWT by an increased number of filters. Al-Otum [100] came up with a robust color-image watermarking for copyright protection, which was based on WPT. Here, the multi-spectral properties of the primary color components of the RGB image are chosen due to their simplicity and high color channel correlation. First, a scrambled watermark is generated by applying the chaos-encryption unit followed by the Arnold transform, which enhances the level of security. Then, after the transformation of the host image using the WPT, the vector information of the transformed image is used to compute the threshold values, which are used to select the embedding place of the watermark. In this sense, the safe locations depend on the inter-layer energy of the wavelet-packets coefficients. In view of the experimental tests, the obtained results reveal a good resistance of the proposed method against a wide range of attacks, with high watermarking imperceptibility. Despite this, it suffers from high computational complexity.

For color image authenticity and copyright protection, Kumar and Singh [101] presented a DWT-based watermarking method using alpha blending and entropy concepts. The DWT is applied on the YCbCr host image to obtain the non-overlapping sub-bands of the "Y" channel. Next, the block with the highest entropy value is adaptively chosen for watermark embedding. Finally, alpha blending is performed to conceal the watermark in the selected block. Results have revealed that alpha blending successfully strikes a balance between the watermarking system's resilience and imperceptibility. The main benefits of the proposed approach include low computational cost, enhanced PSNR value, and high robustness against various sorts of attacks. However, the watermarking scheme requires the host image in the extraction stage and it has only been tested with one kind of watermark data, and with inadequate time and security analysis.

On the other hand, Anand and Singh [102] proposed a Paillier cryptosystem and turbo code-based DWT image watermarking technique for grayscale medical images in telehealth applications. In order to provide authentication, the watermark is first scrambled using a step space-filling curve, then encoded using turbo code. Furthermore, the retrieved watermark is concealed in the host image using the homomorphic properties of the Paillier Cryptosystem. The experimental results show that the discussed technique is very robust and meets all watermarking requirements for telemedicine applications in terms of imperceptibility, robustness, capacity, and security. Since the encryption operation provides a high level of computational complexity, selective encryption techniques can be used to improve performances in this aspect.

Lastly, Yuan et al. [103] designed a secure and blind watermarking method based on discrete wavelet transform and computational ghost imaging (CGI). The watermark is encrypted by the CGI system with a chaotic key. Then, the quantization algorithm is applied to the ciphertext, which compresses the amount of information to be concealed in the wavelet coefficients of the original image. In this approach, the chaotic key is applied in both the embedding and encryption stages, which improves the security of the watermarking scheme. Moreover, it can withstand both salt and pepper noise and Gaussian noise at the same time. However, It should be noted that with increasing noise and cropping, the quality of the reconstructed image gradually declines, and the compression of the ciphertext can induce serious distortion in the retrieved watermark.

The wavelet transform-based image watermarking schemes provide good performance in terms of invisibility and robustness to various types of attacks. Moreover, The main advantage of DWT over DFT is the temporal resolution, which makes DWT a more attractive research area. Yet, it is still sensitive to geometric transformations and it requires a high complexity value.

Table 2 summarizes an overview of the transform domain-based state-of-the-art schemes.

Table 2. Summary of transform domain-based image watermarking techniques.

Reference	Approaches	Objectives	Cover/ Watermark Size	Evaluation Metric	Advantages and Weaknesses
[83]	DCT DE KELM	Achieve good balance between robustness and invisibility	512 × 512 32 × 32	PSNR, NC	+ Robust against numerous attacks - Fragile to rotational attacks
[85]	DCT JND Arnold transform	Achieve good balance between robustness and invisibility	512 × 512 64 × 64	PSNR, SSIM, NC, BER	+ Resist different attacks, such as median filter and image rotation - Security and complexity analysis are inadequate
[94]	QDFT EPA MSPSAM MM, PSO	Provide better robustness against JPEG attacks and contrast enhancement	512 × 512 64 × 64	Objective Value PSNR MSSIM BER	+ Resist several image-processing attacks, including JPEG attacks - High computational time
[95]	DFT, PSO	Improve performance, in particular the security	COCO dataset 32-bit binary watermark	VIF PSNR Average BCR BER	+ Withstand image-processing attacks and geometrical ones - Payload needs to be improved
[101]	DWT Entropy Alpha blending	Reduce computational cost considering robustness and imperceptibility	512 × 512 64 × 64	PSNR SSIM NCC	+ Low computational cost + Higher PSNR value + Robust against various sorts of attacks - Non-blind - Testing with one type of watermark data is inadequate - Security and complexity analysis need to be discussed in detail
[102]	DWT Paillier cryptosystem Turbo code Step space-filling curve	Provide robustness and security of the EPR data	512 × 512 text watermark and different sizes of watermark image	NPSR UACI BER NC	+ Robust and secure for medical-image watermarking - Lack of computational complexity test

4.3. Hybrid Domain

The spatial domain achieves better performance in terms of invisibility, while the transform domain provides high robustness. However, the watermarking schemes proposed in one domain are limited in several scenarios. Watermarking algorithms based on multi-domains are known as hybrids, and these algorithms ensure better robustness and improved information embedding properties. The next part introduces recent approaches in the hybrid domain.

Darwish et al. [104] proposed dual color-image watermarking using DWT, WHT, and SVD to ensure image security for copyright protection. The cover image is converted to the YCbCr color space and the WHT (Hadamard transform) is employed to encrypt the watermarks. Then, the salient features are extracted based on DWT followed by SVD to calculate the singular matrix of the selected sub-band of the DWT image. In order to satisfy the perceptibility when dual watermarks are embedded, the genetic algorithm is adopted to determine the optimal embedding positions and scaling factor. Although the method shows better robustness against common image manipulation attacks, the model is partially blind, requiring more memory for the recovery process, and the time complexity is too high.

The approach in [105] is proposed in the DWT-DCT-SVD domain, wherein the host image is transformed using the DWT to produce LH2 and HL2 sub-bands, and the DCT

is applied to these sub-bands. Finally, the chaotic logistic map is adopted to shuffle the watermark image to be inserted in the SVD domain. Additionally, an optimized version of PSO is proposed to select the most appropriate DCT sub-bands via a multi-dimensional optimization and to estimate the watermark-embedding strength considering the imperceptibility and the robustness of the watermarking scheme, which performs better in terms of watermarking requirements, especially in terms of security. However, some geometrical attacks, such as rotation and scaling, have not been treated.

In [106], a blind medical-image watermarking scheme was designed using hybrid transforms, DWT, and Schur decomposition. In order to achieve both tamper recognition and authenticity, dual watermarks were concealed in the region of the non-interest (RONI) blocks of the images. To determine the embedding position, the discrete wavelet transform and the Schur decomposition are then performed. Both watermarks are compressed by Lempel–Ziv–Welch (LZW) to increase the payload capacity according to image quality, while the swarm bacterial foraging optimization algorithm (PSBFO) is used to achieve the balance between imperceptibility and robustness. Evaluations carried out demonstrated the performance of the watermarking algorithm in terms of imperceptibility and robustness against various attacks. However, since the scheme consists of several mechanisms, as the file size increases, so does the complexity.

For the same purpose, Alzahrani et al. [107] presented a hybrid watermarking scheme intended for the medical field that relies on three collective strategies, including DWT, DCT, and SVD. The region of interest (ROI) and region of non-interest (RONI) are generated from image separation, and the DWT is applied to RONI to produce low- and high-level bands. After the selection of embedding potential blocks using the human visual system (HVS), and the subdivision of these blocks to carrier matrices, the watermark is inserted by modifying the largest diagonal singular values of these matrices. The technique is blind and achieves higher imperceptibility, as well as robustness. However, the combination of these three transformations increases the computational complexity and the risk of false positives.

In [108], a robust rotation-invariant watermarking technique is suggested. In this approach, DWT is performed on the original image, and the DWT-LL sub-band was chosen for embedding. After dividing the LL sub-band into non-overlapping blocks, DCT is performed on random blocks. Then, the singular value matrix of the watermark is concealed in the DCT coefficients. Hence, using singular value decomposition improves the robustness of the watermarking scheme against rotation attacks, while the use of DWT and DCT transforms help to strengthen robustness against common image-processing attacks, such as noise, blurring, and compression. In this case, the security of the system was not considered since no encryption techniques were used to encrypt the watermark.

Recently, Ernawan et al. [109] proposed an improved image watermarking method based on DWT-DCT coefficients. The image blocks with the highest variance values are selected for watermark embedding, which is less sensitive to the human eyes. DWT is applied on these blocks to produce a DWT LL sub-band followed by a two-dimensional DCT. Afterward, a set of embedding rules are used to conceal the watermark, considering the adaptive scaling factor. Additionally, the x and y coordinates of the selected image blocks are adopted for extracting references. In this work, the imperceptibility of the watermark image was verified and the robustness against various attacks was proved, including compression, Gaussian filter, and sharpening. However, the quality of the watermarked image and the robustness against added noise can be improved.

The main idea of the hybrid domain is to combine the characteristics of several domains and adopt them in one method to improve performance and reduce the weakness of the watermarking schemes. On the other hand, this methodology increases computational complexity in some cases.

Table 3 summarizes an overview of the hybrid domain-based state-of-the-art schemes.

Table 3. Summary of hybrid domain-based image watermarking techniques.

Reference	Approaches	Objectives	Cover/ Watermark Size	Evaluation Metric	Advantages and Weaknesses
[108]	DWT DCT SVD	Robustness against rotation attacks	512 × 512 Multiple sizes	NC, PSNR	+ Robust against various types of attacks including rotation attacks - Few attack analytics - Security was not considered in the scheme
[109]	DWT DCT Arnold transform	Improve imperceptibility considering robustness	512 × 512 32 ×32	ARE SSIM PSNR NC BER	+ High PSNR and SSIM values - Lower BER values under various attacks - Limited embedding capacity

5. Deep Learning-Based Image Watermarking Schemes

Previously, various conventional image watermarking methods were described in different domains using different embedding and extraction techniques. These approaches rely on understanding many different types of carrier images and need the careful consideration of several factors in each case when implementing the watermarking scheme. At the present, with the growth of multimedia technologies, using images has become easier but more difficult to process. This makes the task of protecting data information more challenging using conventional methods, thus renovation is needed to speed up the development in this area.

On the other hand, deep learning (DL) is a subset of artificial intelligence based on neural networks to simulate the behavior of the human brain in terms of data processing and decision making [110,111]. DL improves prediction performance using big data and plentiful computing resources. It is used mostly in the image-processing field to solve difficult problems, such as image classification, semantic and instance segmentation, and object detection. Recently, deep learning-based watermarking methods have received extensive attention and have become the state of the art in the image watermarking field [18]. Instead of using an engineered algorithm, DL has the ability to learn complicated features from images to represent labels automatically as a minimization of some loss function on a training set of data. Convolutional neural networks (CNNs) are the most successful subfield of DL-based image-processing techniques, as they underlie many deep learning architectures of data images [112]. Further, they are able to take advantage of graphics processing units (GPUs) for computation, while many other network architectures cannot. In our case, various watermarking methods have been proposed in the literature related to the powerful ability of deep learning [21], especially CNNs, for data hiding and data extraction, which improved automation and carrying out analytical tasks without the need for human intervention. Thus, they provide satisfactory performance and effectively maintain the trade-off between robustness and image quality.

In this section, we will explain the impact of deep learning on the development of the image watermarking field. For this purpose, a comprehensive description of recent learning-based methods is explained in detail. The reviewed methods will be separated into three categories: those that use deep learning in embedding, in extraction, and in both embedding and extraction. Figure 9 describes this classification.

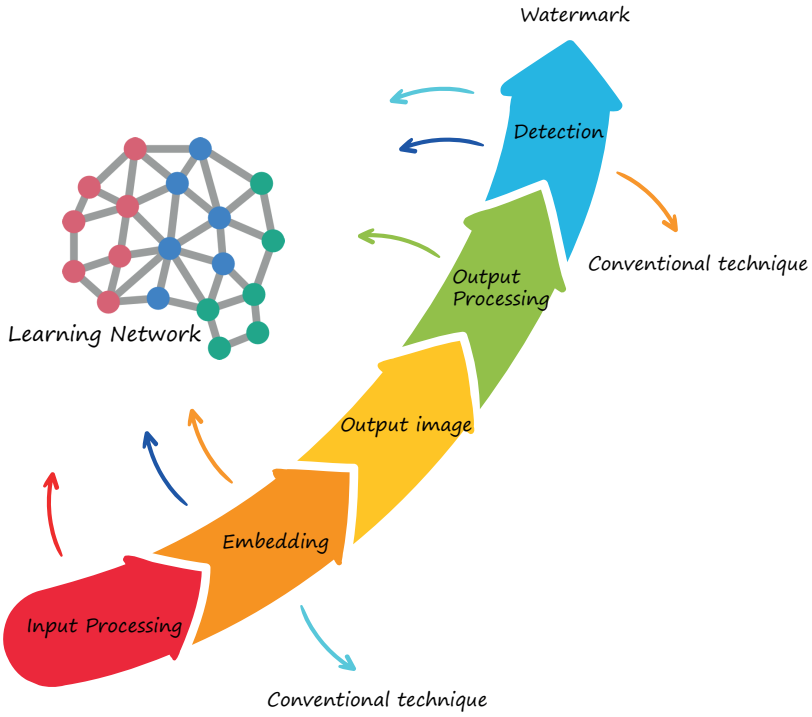


Figure 9. Brief diagram of the learning framework in the image watermarking process. The network was trained either to insert the watermark (orange line) or to detect the watermark (sky blue line). The scenario in which both embedding and detection are performed by a trained network is represented by the dark blue line, while input processing and output processing using learning networks are shown by the red line and the green line, respectively.

5.1. Learning-Based Embedding Techniques

The main objective of the embedding stage is to ensure watermark imperceptibility while also taking into account the other requirements, such as robustness and capacity. For this, many neural networks-based approaches have been designed either to conceal the watermark in the host image or to enhance the watermark-embedding system. As an example, Bagheri et al. [113] proposed a deep learning method to find the appropriate strength factor for watermark embedding in the DWT-DCT domain. In order to calculate the strength factor, the cover image is first fed into the network, which is called R-CNN, to extract masks M_i. The maximum importance coefficient c_i is used to compute the strength factor as follows:

$$M_E = \alpha(1 - \max_i(c_i \times M_i)), i \in \{1, 2, \ldots, k\} \quad (13)$$

where k represents the total number of important classes generated by R-CNN and α is a constant to scale the strength factor M_E. In parallel, The host image is split into non-overlapping blocks and the ROI is assigned to all blocks. Then, DWT is applied to these blocks, followed by the DCT, and one bit is embedded for each sub-block. After that, the strength factor is used to maintain the inserted bits in case the host image has been attacked. Finally, DWT and DCT are used to extract hidden watermarks regarding the 15-times redundancy of the watermark bit applied in the embedding phase. The proposed approach is more imperceptible and resistant against image-processing attacks. However, the robustness against various attacks should be investigated. Moreover, a comparison with a single existing method is insufficient.

In [114], the improved Arnold transform is executed on the watermark image, and the carrier image is compressed using the backpropagation (BP) neural network, producing an output in the hidden layer. Then, both scrambled watermark and watermark normalization processing are applied to the output of the hidden layer to generate the compressed wa-

termark image. In Addition, normalization processing tries to reduce the error prediction of the BP neural network, hence improving prediction accuracy. Finally, the compressed watermarked image is further decompressed using the presented network to obtain the watermarked carrier image. For extracting the watermark, The anti-normalization processing of the difference between the original output and the new output of the hidden layer is calculated. Results reveal the feasibility of the proposed algorithm in terms of imperceptibility, complexity, and robustness against conventional and geometrical attacks. Nevertheless, it shows a poor PSNR value that is less than 20dB for the majority of testing attacks.

On the other hand, Cu et al. [115] suggested a watermarking scheme for image documents using fully connected networks (FCN). First, the input image is standardized to enhance robustness against geometric distortions, and the watermark is further encrypted as a security feature. Next, FCN is used to detect the watermarking regions of the image document. Here, constructed hidden patterns are adopted to find the appropriate positions inside these watermarking regions. The pixel values of these patterns are changed to conceal the secret bits based on the corner feature and the edge of document content and according to their neighboring pixels. Experiment analysis has demonstrated a high stability degree of the FCN against document distortions. Moreover, common image watermarking requirements have been satisfied. However, the proposed scheme still vulnerable to Gaussian filtering, especially printing and scanning attacks at lower resolutions.

With the same motivation, Ingaleshwar et al. [116] designed a deep convolutional neural network associated with water wave optimization (WWO) and the chaotic fruit fly optimization algorithm (CFOA). The network is composed of three layers, including a fully connected layer, and it uses feature input to find the optimal region for embedding the secret message. In addition, the classifier is trained using the water chaotic fruit fly optimization algorithm (WCFOA), which was developed by integrating the WWO and CFOA. To compute the best solution with the minimal error value, a fitness function is calculated as:

$$F = \frac{1}{\sigma} \sum_{v=1}^{\sigma} \beta_{p(v)}^{q} - \overline{\omega}_v \tag{14}$$

where, σ represents the total number of samples, $\overline{\omega}_v$ is the estimated output, and $\beta_{p(v)}^{q}$ represents the output of the classifier. Finally, the watermark is inserted in the wavelet transform sub-bands according to the optimal region and based on the fitness function F, wherein the optimal region helps to effectively increase the embedding performance. At the extraction stage, the cover image is required for extracting the hidden information based on wavelet transform, as the prime objective of the proposed scheme is to achieve a good imperceptibility of the watermarked image. The method provides high values in terms of the correlation coefficient and PSNR. However, working on scheme robustness will increase the scientific value of the present work. Moreover, the computational cost of the scheme must be investigated because various mechanisms have been involved.

In contrast, Sinhal et al. [117] developed an artificial neural network to reduce the computational cost of the embedding process. First, the Y channel of the YCbCr image is divided into blocks.The embedding blocks are then selected in a randomized manner using the Mersenne twister random number generator with a secret key to improve the robustness. Then, the watermark bits are concealed in the IWT-DCT sub-bands via a trained artificial neural network (ANN). Specifically, the selected blocks and four copies of the watermark bit are used to generate 20 values as the input of the ANN, thus modifying the pixel values of a block, while the watermark is extracted by applying IWT on the blocks of the watermarked image followed by using the DCT transform on the LL band. The performance comparison reveals good values in terms of imperceptibility and robustness against different image-processing attacks, as well as fewer values against some of them, including Gaussian noise, JPEG, and median filter.

Further, Kandi et al. [118] designed a novel learning scheme based on two convolutional neural network (CNN) auto-encoders for non-blind image watermarking. The

positive and the negative of the original image are fed to the CNN's auto-encoders, generating two sets of codebook images. These codebook images are permuted and then vectorized. Finally, the watermarked image is obtained from one of the two vectorized codebook images according to the permuted watermark. The use of CNN provides higher security as the weights are initialized randomly, and it learns features of the input to generate two codebook images. In the process of extraction, codebook vectors are generated using the parameters of the trained network in the embedding phase along with keys. The permuted extracted watermark is obtained by computing the Euclidean distance (nearest vector) of the watermarked image from the codebook vectors. Although, the proposed scheme required the original signal and keys for the extraction process, which is not suitable for many real-time applications, the training of the CNN was performed offline, and it is robust against various types of attacks.

Finally, the ability of deep learning classification for blind image watermarking is discussed in [119,120]. These studies show that classification networks produce high-quality images and a good balance between invariance to geometric attacks and robustness to volumetric attacks. Therefore, the features extracted from these networks are flexible and relevant. Moreover, augmenting the data with specific image transformations increases the invariance of geometric transformations. On the other hand, deep learning-based watermarking schemes are not completely invariant to this type of attack. This is suitable for applications that face small geometric transformations; however, a registration system must be used for any other applications.

Table 4 shows some of the watermarking schemes based on learning networks for the embedding process.

Table 4. Summary of learning-based image watermarking methods for embedding process.

Reference	Architecture	Goal	Embedding Location	Cover/Watermark Size	Evaluation Metric	Advantages and Weaknesses
[113]	R-CNN	Find the good stength factor used to determine the appropriate location for embedding	DCT-DWT sub-blocks	512 × 512 4 × 4	PSNR = 49.10 dB SSIM = 0.99 NC = 1 BER = 0	+ High image quality + Robust against several image-processing attacks, such as JPEG, Gaussian noise, and median filter - Comparison with one existing method is inadequate - Robustness should be investigated in detail - Lower embedding capacity
[116]	CNNs	Find the optimal region to conceal the watermark	DWT coefficients	MRI scans	correlation coefficient = 1 PSNR = 45.2 dB (without noise scenario)	+ High values of correlation coefficient and PSNR - Non-blind - Focus only on imperceptibility
[117]	ANN	Reduce computational cost of the embedding process	Y channel coefficients	512 × 512 32 × 32	PSNR = 39.9 dB SSIM = 0.99 (Lena image) Avg BER = 0.05 Avg embedding time = 0.41s	+ Low computational complexity + Robust against different signal-processing operations - Performance should be studied with multiple watermarks - Geometric attacks must be studied
[118]	Auto-encoder	Learning codebook images	Vectorised codebook image	128 × 128 64 × 64	Avg PSNR = 42.03 dB NC = 0.96 crop(25%)	+ High security + Imperceptibility and robustness - Non-blind - High computational cost

5.2. Learning-Based Extraction Techniques

Instead of using the trained neural network for the embedding process, it can be adopted at the extraction stage, either to completely extract the watermark or to improve extraction results. Various contributions-based learning networks were proposed for this purpose. For instance, Li et al. [17] presented a novel CNN-based security-guaranteed image watermarking for smart city applications. The main object of using CNN in this case is to classify images according to the method of image generation. First, a gray watermark image was embedded into the DCT-block component. Then, a cooperative neural network

called Fast R-CNN is used for watermark extraction, which takes the suspected watermark as input and produces an output based on the recognition process. The proposed Fast R-CNN has improved the processing speed. However, the computation of potential regions continues to experience delays; therefore, the model should be investigated using different image databases.

Kazemi et al. [121] designed a neural network-based watermarking framework to deal with color images under a set of attacks. First, The contourlet transform is chosen to produce the directional coefficients, and the Zenzo edge detector generates the edge of the carrier image. The watermark logo is inserted into the edge of the color image appropriately using a genetic algorithm. Finally, differential and multi-layer perceptron (MLP) is used to extract the logo information. Experiments with different scenarios were carried out to verify the effectiveness of the discussed approach. Accordingly, it shows good performance values to deal with different kinds of attacks, even if the method is less powerful against a few attacks, such as histogram equalization and average filtering.

In [122], the authors designed a new image watermarking method based on the combination of DWT, DCT, SVD, and backpropagation neural network (BPNN) for healthcare applications. Third-level DWT, DCT, and SVD are applied on the cover image, and three watermarks are used for embedding including a Lump image watermark, a symptom, and a text watermark. These watermarks are concealed on the DWT sub-band level, which offers better performance in terms of imperceptibility, robustness, and capacity. To improve the security of the scheme, Arnold transforms, lossless arithmetic compression technique, and Hamming error correction code are performed on the three watermarks, respectively, before embedding. After watermarks extraction, BPNN is then applied to the extracted watermarks to remove the noise, therefore, enhancing the robustness. The experimental results indicate that the proposed method is able to resist different image-processing attacks considering robustness, imperceptibility, capacity, and security simultaneously. However, the computational complexity should be investigated due to the combination of many watermarking mechanisms.

Further, Zear et al. [123] came up with a hybrid image watermarking based on discrete wavelet transforms (DWT) and singular value decomposition (SVD) using a backpropagation neural network (BPNN). First, the color image is divided into third-level DWT; then, SVD is used to transform both the low-frequency band LL3 and the watermark image. The S vector of the host image is adopted to conceal the S vector of the watermark information. Finally, the hidden watermark is extracted using an extraction algorithm. BPNN, in this case, is performed on the extracted watermark to reduce the noise effects so as to improve the robustness of the watermark image. The experimental results showed that the presented watermarking scheme is robust against common signal processing, with good imperceptibility, which was the main objective of this work. Although, the combination of many techniques improved the robustness and the imperceptibility, it may also slightly raise processing cost.

In another paper, Huynh-The et al. [124] proposed a blind image watermarking technique that exploited a deep convolutional encoder–decoder network for extraction purposes. The watermark is inserted into selective wavelet coefficients for image imperceptibility enhancement. Then, the embedding maps, defined as the difference values between wavelet coefficients from different attacking simulations of the watermarked image, are used to train the deep learning model for watermark extraction, wherein the trained model can precisely recover the watermark data from its host image. The predicted watermark bit W' corresponds to the maximum probability of the output p of the softmax layer in the network as follows:

$$W' = \underset{i}{\operatorname{argmax}}(p) \tag{15}$$

From the results, this approach achieved a good trade-off between watermark robustness and image imperceptibility. However, encryption techniques are not included; hence, security performance should be investigated.

Likewise, in [125], the watermark is concealed in the LWT-Block to maintain a good balance between imperceptibility and robustness. In the extraction phase, the attacked image is received and then transformed using three-level LWT to extract the sub-band LH/LH1/HL2. The extracted HL2 sub-band is converted into a feature set, which is fed to the trained network to generate the watermark array. The predicted values were then transformed into a matrix to create the watermark. To create a feature vector, the watermarked images from a total of 100 datasets used to train the network were attacked with different types of attacks, including noises, filters, cropping, and scaling. In addition, the prime goal of DNN is to identify the changes made by these attacks to enhance the robustness of the scheme. The proposed study shows a good withstanding of several types of both signal-processing attacks and geometric attacks. However, it does not provide sufficient results for robustness improvements against some image-processing attacks, such as speckle noise and salt and pepper noise. Moreover, rotation and translation issues must be investigated.

The proposed approach in [126] combines a user-specific watermark and a messenger application-specific watermark to form the source watermark. Furthermore, randomized orthogonal codes are used on messenger-specific watermark for better security, and an optimization approach has also been used to reduce the computational execution times for embedding and extraction processes. In order to achieve better reliability, three copies of the source watermark are used for embedding on the Slantlet domain. A multi-layer backpropagation neural network is used to extract the watermark appropriately and quickly, under the pressure of various attacks. In general, the presented method was tested against a variety of signal-processing attacks and it showed high robustness with significant imperceptibility. However, geometric attacks are not studied in this approach. Moreover, the validity of the scheme for different types of watermarks could be the future work.

Lastly, Wang et al. [23] designed a blind extraction (non-embedding) image watermarking framework based on a residual convolution neural network (RCNN) to perform the mapping relationship between the host image and the watermark image. First, the host image is processed using the median filter to improve the robustness and reduce texture information. Then, DCT and SVD were applied to the image sub-blocks to generate the information matrix, which is used as the input of the RCNN. After training, the trained network parameters are securely stored in order to be used in watermark extraction. This network contains several residual blocks, with four convolution filter layers in each residual block, and used an improved cross-entropy loss function to maintain the balance between security and the recognition accuracy of the original images and non-original images. Moreover, random noise is added during the training process to enhance both robustness and the training speed. In order to outperform the false positive detection problem produced by using SVD, the proposed framework supports a multi-type watermark with high-performance security and achieves better results against several types of attacks, especially image-processing attacks. However, it is still less robust to rotation and cropping; additionally, there is a lack of complexity analysis.

Table 5 shows some of the watermarking schemes based on learning networks for the extraction process.

Table 5. Summary of learning-based image watermarking methods for extraction process

Reference	Architecture	Goal	Embedding Location	Cover/Watermark Size	Evaluation Metric	Advantages and Weaknesses
[17]	Fast R-CNN	Learn to extract the watermark	Block-DCT component	- -	Avg. PSNR = 49.10 dB Extraction time = 1.19 s	+ High PSNR values - Focus only on imperceptibility
[123]	BPNN	Remove noise spikes from the watermarked image	S vector of the host image	512 × 512 64 × 64	NC = 0.98 PSNR = 34.78 dB (peppers image, gain = 0.1)	+ Good imperceptibility and robustness - High computational complexity - Performance needs to be investigated with multiple watermaks

Table 5. *Cont.*

Reference	Architecture	Goal	Embedding Location	Cover/Watermark Size	Evaluation Metric	Advantages and Weaknesses
[124]	Encoder–Decoder	Recover robust watermark from the host image	DWT coefficients	512 × 512 64 × 64	Avg NC = 0.99 Avg PSNR = 47.85 dB Avg SSIM = 0.99	+ Withstand image transformations and geometric distortions - High PSNR values - Security performance should be investigated
[125]	DNN	Identify changes made by attacks	Block LWT component	512 × 512 32 × 32	Avg NC = 0.98 Avg PSNR = 44.11 dB	+ Robust against common image-processing attacks + Fast watermark extraction - Needs further improvement in robustness performance
[23]	R-CNN	Map the relationship between the host image and the watermark image	DCT-SVD sub-blocks	512 × 512 32 × 32	NC = 0.9930 (salt and pepper 0.5)	+ Outperform the false detection problem + Support multi-type watermark + High security and robustness - Less robust to rotation and cropping - Lack of complexity analysis

5.3. End-to-End Learning-Based Watermarking Techniques

To date, the previous two categories only focus on one stage of the watermarking system. In contrast, there are many watermarking methods that embed and extract the watermark with the help of neural networks. Here, the learning framework aims to improve performance in both the embedding and extraction phases.

Firstly, Zhong et al. [127] proposed a robust and blind image watermarking based on the mapping ability of deep neural networks. The designed framework was trained in an unsupervised manner to generalize both watermark embedding and extracting processes. It consists of five neural network components trained as a single deep neural network. The first is applied to encode the watermark image, which brings randomness and redundancy to enhance information protection and robustness. The second is called the embedder, it takes the feature space of the cover image with the encoded watermark and produces the watermarked image in a fusion way. Before watermark extraction, another network is adopted called the invariance layer, which converts the watermarked image to its redundant transformed space, wherein this layer rejects irrelevant information about the watermark and preserves the most important information, thus providing a high tolerance of errors on the watermarked image, which enhances robustness. Finally, two neural network components are used to fit the watermark reconstruction, the first extracts the watermark feature space from the output of the invariance layer, and the second decodes the watermark. The reconstruction is blind since it requires only the watermarked image without the need for the watermark and the original image. Moreover, the recovery ability of auto-encoders with appropriate features secures the feasibility of the watermark extraction in the proposed scheme. Experimental results reveal the system's good performance against typical attacks and its applicability in challenging camera applications. However, image fusion slows down data processing and results in huge data losses. Furthermore, geometric attacks should be studied.

In order to deal with rotation attacks and JPEG compression, a blind learning network was developed in [128], which consists of a stego-image generator, an attack simulator, a watermark extractor, and a stego-image discriminator. The generator network learns to embed the watermark image into the cover image, and during the training, a generative adversarial network (GAN) is used to improve the quality of the stego-image. The attack simulator consists of two layers, namely a rotation layer to simulate the rotation attack and an additive layer to simulate the JPEG attack. Therefore, the watermark extractor can extract watermarks without rotation synchronization since it receives rotated and noised images in the training phase. In addition, the weight parameters and the additive noise's strength could be used to adjust the robustness and the image quality of the present scheme. The evaluation test revealed strong resistance to rotation and JPEG compression. However, an attack simulator with a variety of attacks can be used.

Some studies rely on deep neural networks to find a robust watermarking domain instead of the frequency domain. In [129], an automated framework was presented based on reinforcement learning for robust watermarking. Three stages were considered in this case, namely watermarking embedding, attack simulation, and weight updating. In the embedding stage, the image block is modified by the network until the network correctly detects the message in the image. Applying this process for all the blocks produces the watermarked image. For that purpose, two methods were proposed. The first is based on backpropagation, and the second is based on an auto-encoder designed only for embedding. Furthermore, an attack simulation and stochastic gradient descent (SGD) are used for robust feature extraction and to correctly capture the message from the host image, respectively. Finally, the weight parameters of the detector network can be used for watermark extraction. The proposed learning network optimized the robustness while considering the high quality of the retrieved images. Consequently, it provided good robustness against different attacks. Due to system generalization ability, it could withstand seen and unseen attacks in the training phase. However, the designed network did not provide a watermark solution for high-definition images.

In contrast, Zhu et al. [130] integrates CNNs with keypoint detection to solve high-definition image watermarking problems. First, various scale-invariant regions in the host image are acquired. Here, the normalized watermark image is fitted in the center of these regions. This manner of insertion ensures locating the watermark even under geometric distortions. Specifically, the Y channels of the selected regions are concatenated with the normalized watermark, and the result is fed to the embedding network to obtain the marked Y channels. Then, the watermarked regions are obtained via the concatenation of the marked Y channels and the original CbCr channels. After that, the original regions are replaced with the watermarked regions to produce the marked high-definition image. For watermark extraction, as in the embedding stage, the embedding regions are segmented on the marked image with the same method, and their Y channels are input to the extraction network, which is called the revealing network, to retrieve the watermark. Together, the revealing network and the embedding network are trained, and as a result, the total loss backpropagation function is defined by combining SSIM, mean square error (MSE) and binary cross-entropy (BCE) as:

$$I = MSE(C, S) + 1 - SSIM(C, S) + BCE(W, W') \tag{16}$$

where C and S represent the input original image and the output watermarked image. The W and W' represent the original watermark and the output of the revealing network. The results show how resistant the suggested approach is to geometric distortions and signal-processing flaws. However, it shows a high complexity for the embedding process.

In [131], a universal image watermarking scheme was introduced using CNN without restrictions on the cover image and watermark information. The proposed system consists of three principal stages, including a pre-processing network, an embedding network, and an extraction network. Here, normalized host image and scrambled watermarks are pre-processed using the first two networks, respectively, and the outputs are concatenated to generate the input of the second network. In addition, the strength-scaling factor is multiplied by the watermark data to achieve the trade-off between invisibility and robustness. After that, the concatenation result is used as the input of the embedding network to produce the watermarked data, which is re-normalized to generate the watermarked image. Finally, the watermark information produced by the extraction network is unscrambled to obtain the final extracted watermark. Two loss functions were used in this work L_{emb} and L_{ext} for embedding and extraction, respectively, as follows:

$$L_{emb} = \lambda_1 L_1 + \lambda_2 L_2 \tag{17}$$

$$L_{ext} = \lambda_3 L_2 \tag{18}$$

where L_1 represents the MSE between the watermarked image and the original image and L_2 is the MSE between the extracted and the original watermark. In these two equations, λ_1, λ_2, and λ_3 are set to the hyper-parameters that control invisibility and robustness. The proposed technique has high robustness against various pixel-value and geometric attacks, as well as good invisibility for the watermark. Additionally, providing complexity analysis may help to improve the scientific value of this work.

The CNN model is further adopted in [132]. In this case, it takes image blocks as input and learns to detect one-bit messages from each block. The image is updated so that the model recognizes the blocks as the message bit. For that purpose, the gradient-descent method (SGD) is adopted. The loss function from the SGD method is used to protect the embedding invisibility, and the message is properly inserted into the block when the loss is close to zero. Finally, the blocks are combined to generate the watermarked image. A set of simulated attacks was applied on the watermarked covers and images to enhance the CNN efficiency in capturing invariant features for various attacks, as well as to use the attacked watermark as the true label for supervised learning in the next stage. In the extraction process, the CNN weights are updated to extract the watermark correctly. Hence, the message bit is extracted from every block of the distorted image and the watermark can be completely extracted based on a threshold value. The performance was improved during model training, including the embedding process, attacks, and extraction. Experiments show robustness against common signal/image operations, such as noise, JPEG compression, Gaussian filtering, and affine transformation. However, the proposed framework provides a uniform local embedding as the traditional methods.

In [133], an end-to-end deep learning network has been used to create a robust and secure image watermarking scheme. Within the transform domain, the embedding layers calculate the watermarking mask/pattern. Then, the residual mask is computed in the spatial domain via the inverse transform to be added to the host image with the weight of a strength factor. The following loss function is used for network end-to-end training:

$$L = \gamma L_1 + (1 - \gamma) L_2 \tag{19}$$

where γ is the loss ratio, and L_1 and L_2 are the embedding and extraction networks' loss functions, respectively. L_1 represents the imperceptibility of the watermarked image, expressed by the SSIM metric, while L_2 reflects the watermark extraction rate and robustness using the binary cross-entropy. In the experimental tests, various simulation attacks are employed as part of the network to govern resistance to specific attacks. The extraction network consists of a copy of the transform layer used in the embedding phase to represent the watermark image in the transform domain, and then other layers learn to extract the watermark. The framework maintains the trade-off between robustness and imperceptibility by adjusting the training and testing parameters of the network. Although the model performs well against known attacks included in the training phase, it may suffer from poor robustness against unknown attacks.

On the other hand, Kim et al. [134] presented a convolutional neural network-based template architecture for recovering watermark synchronization. The block-based watermark is concealed in the image blocks through the watermark-embedder network. Then, the template generation network produces a noisy template to be added to the image. The extraction network finds spatial locations where the template is inserted in the image. The resulting resized and original watermarks are input to the template-matching network to estimate the RST parameters, which are used to further recover the image with geometric distortions as a step for image synchronization. Finally, the watermark decoder is adopted to extract the watermark based on the block size and the position information used in the embedding step. The proposed scheme exhibits good resistance to strong geometric and simultaneous attacks. However, the processing complexity of the template-matching network increases as robustness improves. In this approach, only half of the image blocks are used to insert information bits; therefore, processing large-scale images may be a weakness.

Similar to the zero watermarking schemes, the proposed network in [135] consists of two stages: the master share generation and image verification, and it learns to recognize two classes: image target and no-image target. The CNN architecture consists of 13 convolutional layers and 2 fully connected layers. First, the inherent image features are extracted through these convolutional layers, and the fully connected layers are used to encode these image features and generate robust ones. These features are then converted into binary data and the secret key is used to permute the watermark pattern. After that, the master share is generated by the XOR operation between the binary data and the permuted binary watermark. Using the same secret key, the watermark pattern is obtained from the extracted permuted watermark pattern retrieved in the same manner as in the sharing phase. From experiments, the proposed scheme is robust against common image-processing attacks, such as JPEG compression and filtering. However, it is still weak to synchronization attacks in terms of robustness and imperceptibility. Moreover, the study does not provide comparison results with existing techniques.

More recently, Ge et al. [136] designed a document-image watermarking scheme for the screen-shooting process using a deep neural network. This framework is an end-to-end neural framework with an encoder to embed the watermark, a distortion layer to simulate the screen-shooting attacks, and a decoder to extract the watermark. In the training stage, the watermark is expanded by a fully connected layer and reshaped to match the size of the cover image; then, it is concatenated with the cover. The result is fed into the encoder network to generate the watermarked image. Next, the distortion layer takes the watermarked image as input to produce the distorted watermarked image. Finally, the decoder network is applied to extract the watermark. During the training and testing phases, adjustment of the embedding strength was proposed to improve the visual quality of the watermarked image. After training, the output of the encoder is directly used as the input of the decoder. Although the captures were taken with care using a camera holder, a high performance in terms of robustness and image quality under screen-shooting distortions was provided.

Table 6 shows some of the watermarking schemes based on learning networks for the both embedding and extraction processes.

Table 6. Summary of learning-based image watermarking methods for embedding and extraction processes.

Reference	Architecture	Goal	Embedding Location	Cover/Watermark Size	Evaluation Metric	Advantages and Weaknesses
[127]	Image fusion framework	Improve robustness and efficiency	Concatenation of the host image with the watermark image	128 × 128 32 × 32	PSNR = 39.72 dB BER = 0 (cropping 20%)	+ Good performance against typical attacks + Require only the watermarked image for extraction + Applicable in challenging camera processes - Image fusion leads to data loss and slows data processing - Geometric attacks must be investigated
[128]	Generator–discriminator network	Achieve robustness against rotation and JPEG compression	Luminosity value of the image	4608 × 3456 8 bit	Avg PSNR = 36.3 dB BER = 0.08 (JPEG Q = 50)	+ High robustness against rotation and JPEG compression attacks - Could resist only a limited number of attacks - Execution time must be investigated - Low embedding capacity
[129]	Backpropagation, Auto-encoders WMnet	Find the robust domain from attacks	Image blocks	512 × 512 24 bit	PSNR = 40 dB NC = 1 (Rotation 10°) SSIM = 0.98 NC = 1 (Gaussian filtering)	+ Optimized robustness and high image quality - Solution for high-definition image is not provided - Low NC values for rotation and cropping

Table 6. *Cont.*

Reference	Architecture	Goal	Embedding Location	Cover/Watermark Size	Evaluation Metric	Advantages and Weaknesses
[130]	Keypoint detection MDResNet	Solve high-definition image watermarking problems	Various scale-invariant regions	512 × 512 px 512 × 512 bit	PSNR = 43.68 dB SSIM = 0.97 NC = 0.96 BER = 0.01	+ Robust to both common signal operations and geometric attacks + Support high-definition image - High embedding execution time - Robustness to several attacks must be improved
[131]	CNNs	Learn invisible and robust watermarking	Concatenation of the pre-processed cover and the pre-processed watermark	128 × 128 8 × 8	PSNR = 40.58 dB BER = 0.01 (JPEG s = 90)	+ Embedding and extraction without restrictions on the cover and the watermark + Good balance between invisibility and robustness - BER values for some geometric attacks are inadequate - Lack of complexity analysis
[132]	CNNs	Secure robustness of watermarking	Host image blocks	512 × 512 64 × 64	PSNR = 39.9 dB (peppers image) NC = 0.98 (Resizing 25%, with registration)	+ Withstand different types of attacks - Higher detection time - Poor results against rotation
[133]	ReDMark	Learn new watermarking algorithm in any desired transform domain	Addition of the residual watermark and the host image	512 × 512 32 × 32	PSNR = 40.24 dB SSIM = 0.98 (α = 0.6) BER = 1.6 (JPEG)	+ Improved security and robustness especially against JPEG - Withstand only known attacks from training phase - Increasing strength factor decreases PSNR and SSIM
[134]	Template generation network	Recover watermark from geometric distortions	Predefined locations of the image	512 × 512 512 × 512	PSNR = 43.66 dB SSIM = 0.98 BER = 0.12 (Rotation 50°)	+ Robust to geometric attacks + High imperceptibility - Large-scale image must be investigated - Focus only on geometric attacks - Lack of execution time analysis
[135]	Master share framework	Learn zero watermarking	Combination of inherent image features with owner's watermark sequence	300 × 300 10 × 10	PSNR = 33.15 dB (Compression 100) BER = 0.01 NC = 0.98 (Gaussian filtering $\sigma = 3$)	+ Low computational cost + Robust against several types of attacks - Lack of comparison with existing techniques - Robustness analysis are inadequate especially for synchronization attacks
[136]	Encoder–Decoder	Resist screen-shooting attacks	Concatenation of the cover and the watermark	400 × 400 100 bit	PSNR = 34.10 dB SSIM = 0.91 CPP = 6.88 Average bit accuracy > 97	+ High performance against screen-shooting attacks - Lower embedding capacity - Testing with one type of watermark data is inadequate - Image captures are taken with care

6. Discussion, Issues and Opportunities

Clearly, the primary goal of watermarking research is to balance the restrictive requirements of imperceptibility, robustness, and embedding capacity. However, it is difficult to maintain a good relationship between those three factors. Most of the methods have improved one property or two but at the expense others. Security and computational complexity are further equally crucial, which increases the difficulty of maintaining a good relationship between all these potential requirements. As a result, several concerns and challenges must be addressed in this field.

Image watermarking research was carried out to exploit various embedding domains. As mentioned previously, different domains have their own advantages depending on specific applications. Several approaches have been proposed in the spatial [57,72], transform [86,92], and hybrid domains [105,106] to maintain the trade-off between watermark properties. The spatial domain saves time and improves efficiency, yet it is vulnerable to most types of attacks, while the transform domain offers high robustness and cost compared with spatial domain-based watermarking. In contrast, the hybrid domain came as a solution to fill in the gap between these domains. However, hybrid methods suffer from a high computational cost, which is not suitable for several real-time applications.

Encryption algorithms were designed in certain approaches from the perspective of watermark security [82,103]. However, encryption raises the overall computational cost of the process, and the security analysis of some methods is insufficient and should

be further investigated in detail. Likewise, the weights of neural networks are usually randomly initialized, thus, the learning ability of these networks provide higher security performance [118].

Many factors greatly affect the embedding capacity, including robustness, image quality level, and cover image size. In this regard, the current studies have suggested some ways to increase the embedding capacity of the data based on different techniques, such as domain properties and compression algorithms [106].

The three types of image watermarking techniques were separated according to the requirements of the extraction process. All suggested solutions are either semi-blind watermarking [83], blind watermarking [23,95], or non-blind watermarking [74,118]. However, the non-blind watermarking methods require the original image to retrieve the hidden information, which is not suitable for certain real-time applications.

In most color-image watermarking schemes, only a single watermark is embedded in the host image. These methods are ineffective in preserving both content integrity and copy protection at the same time. Unfortunately, only a few dual watermarking approaches have been proposed to address this issue [81,104].

The watermark preparation step may include watermark scrambling using a chaotic map or Arnold transform [105,122], which enhances robustness against cropping attacks and added noise. However, they are less helpful for attacks that alter the entire image, such as rotation, scaling, and lighting changes.

In order to meet the watermarking requirements, various watermarking strategies have used optimization techniques to obtain the appropriate embedding locations or the optimal strength factor [95]. Additionally, some approaches suggested neural networks as an alternative to these optimization algorithms [113,116]. Nevertheless, the complexity cost of these methods is high.

Usually, a robust embedding space is obtained by using an invariant domain or by combining the basic domains; however, today, neural networks can learn robust watermarking domains for different situations [129]. The gain in robustness, however, is always at the expense of other watermarking requirements.

Many studies focus primarily on resisting one type of attack [134], which is insufficient for certain applications. Further, some approaches focus on one watermarking requirement while ignoring the others [17].

On the other hand, most watermarking schemes are susceptible to synchronization attacks. Various attempts have been made to increase robustness against this type of attack; however, only a few schemes were tested in real-world conditions [90,127].

Several schemes have used neural networks for embedding or extraction processes; however, they are frequently used as an additional step or complementary procedure. Some have adopted BBPN in the extraction stage to remove the noise effects, such as robustness improvements [123]. However, BPNN is inappropriate for processing huge amounts of data since BPNN depends on connections between nearby pixels, while certain studies have used neural networks to learn image fusion or concatenation capabilities as a solution to enhance robustness and imperceptibility [127,130]. Nevertheless, this resulted in huge data losses and high computational time.

Designing a deep neural network model needs a large amount of data to train. Unfortunately, even with augmentation techniques, collecting and preparing suitable datasets in large quantities is difficult, especially for real-world applications. In addition, small training datasets have a significant negative impact on model efficiency, which can be another issue for industrial applications. Usually, a learning model is trained using a particular type of dataset [136] and thus may not be applicable to another type of dataset.

In addition, many researchers have used attack simulation networks to improve the robustness of the watermarking schemes [128,133]. The original image is therefore exposed to simulated attacks during the training phase, called seen attacks, which increases the ability of the scheme to deal with these attacks. However, the trained model can withstand only a limited number of attacks, and may not be effective against unseen attacks.

In general, learning-based watermarking schemes also aim to find a balance between robustness and imperceptibility. However, these techniques focus only on differentiable transformations, making them vulnerable to removal attacks. Retraining the model with clean datasets, for example, can remove unstable watermarks [137]. The only solution for this kind of issue is to increase the performance of the model as much as possible. Furthermore, attackers can use fake images to destroy the watermark. In fact, modern computer vision techniques can transform the entirety of the image content. Therefore, training the deepfake generation networks with host images to produce fake ones may remove the watermark. Several efforts have been made to prevent or disturb the reuse of the watermarked image [138,139].

Image watermarking techniques often used RGB and YCbCr color spaces. The RGB model consists of highly correlated R, G, and B channels, which offer higher capacity and correlation for the watermarking scheme. Likewise, The YCbCr color space was introduced to be more appropriate for hiding secret information, specifically in the Y channel. Despite the useful correlation between RGB channels, it is still not sufficiently used by neural networks, resulting in feature information loss. Moreover, the neural network models that are trained to learn features in one color space exhibit artifacts in the other color space [140]. Therefore, additional research should be conducted on the color space issue, since it might be used by a malicious user to discover the existence of the watermark.

To summarize, the field of image watermarking has seen amazing growth as a result of various contributions produced to protect image content, particularly since the emergence of deep learning, which has substantially accelerated the field's progress and openness to modern technology. However, the traditional watermarking field is very mature and, in our opinion (shared by the community), it has reached its limits. Using deep learning for watermarking is relatively novel. Therefore, there are opportunities to find more relevant schemes to provide the best "defense" against "attacks".

7. Conclusions

In this paper, we provided a comprehensive review of digital image watermarking regarding conventional techniques and learning-based techniques. First, we discussed the background of the image watermarking domain, including the watermarking phases and different kinds of attacks. Second, the image watermarking requirements and their corresponding evaluation metrics were described. We then reviewed and summarized some of the new conventional image watermarking schemes in detail, including the approaches used, their objectives, advantages, and weaknesses. Next, we presented another new comprehensive review of learning-based watermarking techniques. These methods were discussed in depth and then summarized as well in detail, including the architectures employed, their objectives, embedding locations, benefits, and drawbacks. Finally, we examined recent challenges and issues in the image watermarking field based on the reviewed methods, as well as some potential solutions. In this regard, we believe that this paper illustrates the transition of image watermarking from relying solely on conventional methods to adopting learning-based techniques, and we hope that this research can help readers understand the conventional and modern techniques in topic. Hence, future work can be expanded by combining traditional and novel learning methodologies to meet the critical needs of watermarking techniques. Furthermore, in order to improve robustness and security, researchers should focus on developing new, advanced methodologies and exploiting the benefits of both modern and conventional techniques.

Author Contributions: Writing—original draft, S.B.; Writing—review & editing, S.B., R.R. and F.R.; Validation, R.R., H.D., F.R. and R.H.; Supervision, R.R., H.D., F.R. and R.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pouyanfar, S.; Yang, Y.; Chen, S.C.; Shyu, M.L.; Iyengar, S. Multimedia big data analytics: A survey. *ACM Comput. Surv. (CSUR)* **2018**, *51*, 1–34. [CrossRef]
2. Sengupta, A.; Mohanty, S. Everything you Wanted to Know About Watermarking: From Paper Marks to Hardware Protection. *IEEE Consum. Electron. Mag.* **2016**, *6*, 83–91. [CrossRef]
3. Agarwal, N.; Singh, A.K.; Singh, P.K. Survey of robust and imperceptible watermarking. *Multimed. Tools Appl.* **2019**, *78*, 8603–8633. [CrossRef]
4. Mishra, R.; Bhanodiya, P. A review on steganography and cryptography. In Proceedings of the 2015 International Conference on Advances in Computer Engineering and Applications, Ghaziabad, India, 19–20 March 2015; pp. 119–122.
5. Evsutin, O.; Melman, A.; Meshcheryakov, R. Digital steganography and watermarking for digital images: A review of current research directions. *IEEE Access* **2020**, *8*, 166589–166611. [CrossRef]
6. Rakhmawati, L.; Wirawan, W.; Suwadi, S. A recent survey of self-embedding fragile watermarking scheme for image authentication with recovery capability. *EURASIP J. Image Video Process.* **2019**, *2019*, 61. [CrossRef]
7. Begum, M.; Uddin, M.S. Digital image watermarking techniques: A review. *Information* **2020**, *11*, 110. [CrossRef]
8. Ahvanooey, M.T.; Li, Q.; Zhu, X.; Alazab, M.; Zhang, J. ANiTW: A novel intelligent text watermarking technique for forensic identification of spurious information on social media. *Comput. Secur.* **2020**, *90*, 101702. [CrossRef]
9. Sharma, C.; Bagga, A.; Singh, B.K.; Shabaz, M. A novel optimized graph-based transform watermarking technique to address security issues in real-time application. *Math. Probl. Eng.* **2021**, *2021*, 5580098. [CrossRef]
10. Mohammed, A.A.; Abdullah, M.A.; Awad, S.R.; Alghareb, F.S. A Novel FDCT-SVD Based Watermarking with Radon Transform for Telemedicine Applications. *Int. J. Intell. Eng. Syst.* **2022**, *15*, 343–354.
11. Sharma, S.; Zou, J.J.; Fang, G. A Novel Multipurpose Watermarking Scheme Capable of Protecting and Authenticating Images With Tamper Detection and Localisation Abilities. *IEEE Access* **2022**, *10*, 85677–85700. [CrossRef]
12. Wazirali, R.; Ahmad, R.; Al-Amayreh, A.; Al-Madi, M.; Khalifeh, A. Secure watermarking schemes and their approaches in the IoT technology: An overview. *Electronics* **2021**, *10*, 1744. [CrossRef]
13. Lee, Y.J.; Na, W.S. E-Passport Advanced Security Technique Using Biometric Information Watermarking. *J. Comput. Theor. Nanosci.* **2021**, *18*, 1540–1549. [CrossRef]
14. Kumar, M.; Aggarwal, J.; Rani, A.; Stephan, T.; Shankar, A.; Mirjalili, S. Secure video communication using firefly optimization and visual cryptography. *Artif. Intell. Rev.* **2022**, *55*, 2997–3017. [CrossRef]
15. Anand, A.; Singh, A.K. Watermarking techniques for medical data authentication: A survey. *Multimed. Tools Appl.* **2021**, *80*, 30165–30197. [CrossRef]
16. Ross, A.; Banerjee, S.; Chowdhury, A. Security in smart cities: A brief review of digital forensic schemes for biometric data. *Pattern Recognit. Lett.* **2020**, *138*, 346–354. [CrossRef]
17. Li, D.; Deng, L.; Gupta, B.B.; Wang, H.; Choi, C. A novel CNN based security guaranteed image watermarking generation scenario for smart city applications. *Inf. Sci.* **2019**, *479*, 432–447. [CrossRef]
18. Wan, W.; Wang, J.; Zhang, Y.; Li, J.; Yu, H.; Sun, J. A comprehensive survey on robust image watermarking. *Neurocomputing* **2022**, *488*, 226–247. [CrossRef]
19. Abdullatif, M.; Zeki, A.M.; Chebil, J.; Gunawan, T.S. Properties of digital image watermarking. In Proceedings of the 2013 IEEE 9th International Colloquium on Signal Processing and Its Applications, Kuala Lumpur, Malaysia, 8–10 March 2013; pp. 235–240.
20. Jiao, L.; Zhao, J. A Survey on the New Generation of Deep Learning in Image Processing. *IEEE Access* **2019**, *7*, 172231–172263. [CrossRef]
21. Amrit, P.; Singh, A.K. Survey on watermarking methods in the artificial intelligence domain and beyond. *Comput. Commun.* **2022**, *188*, 52–65. [CrossRef]
22. Sy, N.C.; Kha, H.H.; Hoang, N.M. An efficient robust blind watermarking method based on convolution neural networks in wavelet transform domain. *Int. J. Mach. Learn. Comput* **2020**, *10*, 675–684. [CrossRef]
23. Wang, X.; Ma, D.; Hu, K.; Hu, J.; Du, L. Mapping based residual convolution neural network for non-embedding and blind image watermarking. *J. Inf. Secur. Appl.* **2021**, *59*, 102820. [CrossRef]
24. Usha Nandini, D.; Divya, S. A literature survey on various watermarking techniques. In Proceedings of the 2017 International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 19–20 January 2017; pp. 1–4. [CrossRef]
25. Rani, B.U.; Praveena, B.; Ramanjaneyulu, K. *Literature Review on Digital Image Watermarking*; ICARCSET '15; Association for Computing Machinery: New York, NY, USA, 2015. [CrossRef]
26. Salah, E.; Amine, K.; Redouane, K.; Fares, K. A Fourier transform based audio watermarking algorithm. *Appl. Acoust.* **2021**, *172*, 107652. [CrossRef]
27. Yu, X.; Wang, C.; Zhou, X. A survey on robust video watermarking algorithms for copyright protection. *Appl. Sci.* **2018**, *8*, 1891. [CrossRef]
28. Kamaruddin, N.S.; Kamsin, A.; Por, L.Y.; Rahman, H. A review of text watermarking: Theory, methods, and applications. *IEEE Access* **2018**, *6*, 8011–8028. [CrossRef]
29. Evsutin, O.; Dzhanashia, K. Watermarking schemes for digital images: Robustness overview. *Signal Process. Image Commun.* **2022**, *100*, 116523. [CrossRef]

30. Hosam, O. Attacking image watermarking and steganography-a survey. *Int. J. Inf. Technol. Comput. Sci.* **2019**, *11*, 23–37. [CrossRef]
31. Cox, I.J.; Kilian, J.; Leighton, F.T.; Shamoon, T. Secure spread spectrum watermarking for multimedia. *IEEE Trans. Image Process.* **1997**, *6*, 1673–1687. [CrossRef]
32. Cox, I.; Miller, M.; Bloom, J.; Fridrich, J.; Kalker, T. *Digital Watermarking and Steganography*; Morgan Kaufmann: Burlington, MA, USA, 2007.
33. Zhang, L.; Wei, D. Robust and reliable image copyright protection scheme using downsampling and block transform in integer wavelet domain. *Digit. Signal Process.* **2020**, *106*, 102805. [CrossRef]
34. Riad, R.; Ros, F.; Harba, R.; Douzi, H.; El Hajji, M. Pre-processing the cover image before embedding improves the watermark detection rate. In Proceedings of the 2014 Second World Conference on Complex Systems (WCCS), Agadir, Morocco, 10–12 November 2014; pp. 705–709. [CrossRef]
35. Kishore, R.R. A novel and efficient blind image watermarking in transform domain. *Procedia Comput. Sci.* **2020**, *167*, 1505–1514.
36. Yuan, Z.; Liu, D.; Zhang, X.; Su, Q. New image blind watermarking method based on two-dimensional discrete cosine transform. *Optik* **2020**, *204*, 164152. [CrossRef]
37. Kaur, M.; Neeru, N. A Review on Digital Watermarking Using LSB. *Int. J.* **2015**, *5*, 210–2014.
38. Chen, B.; Wornell, G.W. Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. *IEEE Trans. Inf. Theory* **2001**, *47*, 1423–1443. [CrossRef]
39. Rosales-Roldan, L.; Chao, J.; Nakano-Miyatake, M.; Perez-Meana, H. Color image ownership protection based on spectral domain watermarking using QR codes and QIM. *Multimed. Tools Appl.* **2018**, *77*, 16031–16052. [CrossRef]
40. Zhang, X.; Su, Q.; Yuan, Z.; Liu, D. An efficient blind color image watermarking algorithm in spatial domain combining discrete Fourier transform. *Optik* **2020**, *219*, 165272. [CrossRef]
41. Anbarjafari, G.; Ozcinar, C. Imperceptible non-blind watermarking and robustness against tone mapping operation attacks for high dynamic range images. *Multimed. Tools Appl.* **2018**, *77*, 24521–24535. [CrossRef]
42. Muñoz-Ramirez, D.O.; Ponomaryov, V.; Reyes-Reyes, R.; Kyrychenko, V.; Pechenin, O.; Totsky, A. A robust watermarking scheme to JPEG compression for embedding a color watermark into digital images. In Proceedings of the 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT), Ukraine, Kyiv, 24–27 May 2018; pp. 619–624.
43. Licks, V.; Jordan, R. Geometric attacks on image watermarking systems. *IEEE Multimed.* **2005**, *12*, 68–78. [CrossRef]
44. Kutter, M.; Petitcolas, F.A. Fair benchmark for image watermarking systems. In *Security and Watermarking of Multimedia Contents*; SPIE: Bellingham, DC, USA, 1999; Volume 3657, pp. 226–239.
45. Kutter, M.; Voloshynovskiy, S.V.; Herrigel, A. Watermark copy attack. In *Security and Watermarking of Multimedia Contents II*; SPIE: Bellingham, DC, USA, 2000; Volume 3971, pp. 371–380.
46. Pal, P.; Singh, H.V.; Verma, S.K. Study on watermarking techniques in digital images. In Proceedings of the 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 11–12 May 2018; pp. 372–376.
47. Fares, K.; Khaldi, A.; Redouane, K.; Salah, E. DCT & DWT based watermarking scheme for medical information security. *Biomed. Signal Process. Control* **2021**, *66*, 102403.
48. Verma, V.S.; Jha, R.K. An overview of robust digital image watermarking. *IETE Tech. Rev.* **2015**, *32*, 479–496. [CrossRef]
49. Qin, C.; Ji, P.; Zhang, X.; Dong, J.; Wang, J. Fragile image watermarking with pixel-wise recovery based on overlapping embedding strategy. *Signal Process.* **2017**, *138*, 280–293. [CrossRef]
50. Zhang, H.; Wang, C.; Zhou, X. Fragile watermarking for image authentication using the characteristic of SVD. *Algorithms* **2017**, *10*, 27. [CrossRef]
51. Feng, B.; Li, X.; Jie, Y.; Guo, C.; Fu, H. A novel semi-fragile digital watermarking scheme for scrambled image authentication and restoration. *Mob. Netw. Appl.* **2020**, *25*, 82–94. [CrossRef]
52. Qi, X.; Xin, X. A singular-value-based semi-fragile watermarking scheme for image content authentication with tamper localization. *J. Vis. Commun. Image Represent.* **2015**, *30*, 312–327. [CrossRef]
53. Zhou, N.R.; Hou, W.M.X.; Wen, R.H.; Zou, W.P. Imperceptible digital watermarking scheme in multiple transform domains. *Multimed. Tools Appl.* **2018**, *77*, 30251–30267. [CrossRef]
54. Cheddad, A.; Condell, J.; Curran, K.; Mc Kevitt, P. Digital image steganography: Survey and analysis of current methods. *Signal Process.* **2010**, *90*, 727–752. [CrossRef]
55. Wang, C.; Zhang, H.; Zhou, X. A self-recovery fragile image watermarking with variable watermark capacity. *Appl. Sci.* **2018**, *8*, 548. [CrossRef]
56. Sutojo, T.; Rachmawanto, E.H.; Sari, C.A. Fast and efficient image watermarking algorithm using discrete tchebichef transform. In Proceedings of the 2017 5th International Conference on Cyber and IT Service Management (CITSM), Denpasar, Bali, Indonesia, 8–10 August 2017; pp. 1–5.
57. Bhalerao, S.; Ansari, I.A.; Kumar, A. A secure image watermarking for tamper detection and localization. *J. Ambient. Intell. Humaniz. Comput.* **2021**, *12*, 1057–1068. [CrossRef]
58. Zhou, N.R.; Luo, A.W.; Zou, W.P. Secure and robust watermark scheme based on multiple transforms and particle swarm optimization algorithm. *Multimed. Tools Appl.* **2019**, *78*, 2507–2523. [CrossRef]
59. Bravo-Solorio, S.; Nandi, A.K. Secure fragile watermarking method for image authentication with improved tampering localisation and self-recovery capabilities. *Signal Process.* **2011**, *91*, 728–739. [CrossRef]

60. Makbol, N.M.; Khoo, B.E.; Rassem, T.H. Security analyses of false positive problem for the SVD-based hybrid digital image watermarking techniques in the wavelet transform domain. *Multimed. Tools Appl.* **2018**, *77*, 26845–26879. [CrossRef]
61. Ansari, I.A.; Pant, M.; Ahn, C.W. Robust and false positive free watermarking in IWT domain using SVD and ABC. *Eng. Appl. Artif. Intell.* **2016**, *49*, 114–125. [CrossRef]
62. Kumar, S.; Dutta, A. Performance analysis of spatial domain digital watermarking techniques. In Proceedings of the 2016 International Conference on Information Communication and Embedded Systems (ICICES), Chennai, India, 25–26 February 2016; pp. 1–4.
63. Pardhu, T.; Perli, B.R. Digital image watermarking in frequency domain. In Proceedings of the 2016 International Conference on Communication and Signal Processing (ICCSP), Tamilnadu, India, 6–8 April 2016; pp. 0208–0211.
64. Begum, M.; Uddin, M.S. Analysis of digital image watermarking techniques through hybrid methods. *Adv. Multimed.* **2020**, *2020*, 7912690. [CrossRef]
65. Tanwar, L.; Panda, J. Review of different transforms used in digital image watermarking. In Proceedings of the 2018 2nd IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India, 22–24 October 2018; pp. 1165–1171.
66. Fazli, S.; Khodaverdi, G. Trade-Off between Imperceptibility and Robustness of LSB Watermarking Using SSIM Quality Metrics. In Proceedings of the 2009 Second International Conference on Machine Vision, Dubai, United Arab Emirates, 28–30 December 2009; pp. 101–104. [CrossRef]
67. Tanaka, K.; Nakamura, Y.; Matsui, K. Embedding secret information into a dithered multi-level image. In Proceedings of the IEEE Conference on Military Communications, Monterey, CA, USA, 30 September–3 October 1990; Volume 1, pp. 216–220. [CrossRef]
68. Yeo, I.K.; Kim, H. Generalized patchwork algorithm for image watermarking. *Multimed. Syst.* **2003**, *9*, 261–265. [CrossRef]
69. Thongkor, K.; Amornraksa, T.; Delp, E.J. Digital watermarking for camera-captured images based on just noticeable distortion and Wiener filtering. *J. Vis. Commun. Image Represent.* **2018**, *53*, 146–160. [CrossRef]
70. Pan-Pan, N.; Xiang-Yang, W.; Yu-Nan, L.; Hong-Ying, Y. A robust color image watermarking using local invariant significant bitplane histogram. *Multimed. Tools Appl.* **2017**, *76*, 3403–3433. [CrossRef]
71. Belferdi, W.; Behloul, A.; Noui, L. A Bayer pattern-based fragile watermarking scheme for color image tamper detection and restoration. *Multidimens. Syst. Signal Process.* **2019**, *30*, 1093–1112. [CrossRef]
72. Abraham, J.; Paul, V. An imperceptible spatial domain color image watermarking scheme. *J. King Saud Univ.-Comput. Inf. Sci.* **2019**, *31*, 125–133. [CrossRef]
73. Gull, S.; Loan, N.A.; Parah, S.A.; Sheikh, J.A.; Bhat, G.M. An efficient watermarking technique for tamper detection and localization of medical images. *J. Ambient. Intell. Humaniz. Comput.* **2020**, *11*, 1799–1808. [CrossRef]
74. Rinki, K.; Verma, P.; Singh, R.K. A novel matrix multiplication based LSB substitution mechanism for data security and authentication. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 5510–5524. [CrossRef]
75. Wang, H.; Su, Q. A color image watermarking method combined QR decomposition and spatial domain. *Multimed. Tools Appl.* **2022**, *81*, 37895–37916. [CrossRef]
76. Mustaqim Abrar, M.; Pal, A.; Sazzad, S. Bit Plane Slicing and Quantization-Based Color Image Watermarking in Spatial Domain. In *Proceedings of International Joint Conference on Advances in Computational Intelligence*; Springer: Berlin, Gernmany, 2021; pp. 371–383.
77. Araghi, T.K.; Manaf, A.; Zamani, M.; Araghi, S.K. A survey on digital image watermarking techniques in spatial and transform domains. *Int. J. Adv. Image Process. Tech.* **2016**, *3*, 6–10.
78. Jimson, N.; Hemachandran, K. DFT Based Coefficient Exchange Digital Image Watermarking. In Proceedings of the 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 14–15 June 2018; pp. 567–571. [CrossRef]
79. Mahto, D.K.; Singh, A.K. A survey of color image watermarking: State-of-the-art and research directions. *Comput. Electr. Eng.* **2021**, *93*, 107255. [CrossRef]
80. Tsui, T.K.; Zhang, X.P.; Androutsos, D. Color image watermarking using multidimensional Fourier transforms. *IEEE Trans. Inf. Forensics Secur.* **2008**, *3*, 16–28. [CrossRef]
81. Rangel-Espinoza, K.; Fragoso-Navarro, E.; Cruz-Ramos, C.; Reyes-Reyes, R.; Nakano-Miyatake, M.; Pérez-Meana, H.M. Adaptive removable visible watermarking technique using dual watermarking for digital color images. *Multimed. Tools Appl.* **2018**, *77*, 13047–13074. [CrossRef]
82. Loan, N.A.; Hurrah, N.N.; Parah, S.A.; Lee, J.W.; Sheikh, J.A.; Bhat, G.M. Secure and robust digital image watermarking using coefficient differencing and chaotic encryption. *IEEE Access* **2018**, *6*, 19876–19897. [CrossRef]
83. Vishwakarma, V.P.; Sisaudia, V. Gray-scale image watermarking based on DE-KELM in DCT domain. *Procedia Comput. Sci.* **2018**, *132*, 1012–1020. [CrossRef]
84. Moosazadeh, M.; Ekbatanifard, G. A new DCT-based robust image watermarking method using teaching-learning-based optimization. *J. Inf. Secur. Appl.* **2019**, *47*, 28–38. [CrossRef]
85. Wang, J.; Wan, W.B.; Li, X.X.; De Sun, J.; Zhang, H.X. Color image watermarking based on orientation diversity and color complexity. *Expert Syst. Appl.* **2020**, *140*, 112868. [CrossRef]
86. Zhang, Y.; Wang, Z.; Zhan, Y.; Meng, L.; Sun, J.; Wan, W. JND-aware robust image watermarking with tri-directional inter-block correlation. *Int. J. Intell. Syst.* **2021**, *36*, 7053–7079. [CrossRef]

87. Liao, X.; Li, K.; Yin, J. Separable data hiding in encrypted image based on compressive sensing and discrete fourier transform. *Multimed. Tools Appl.* **2017**, *76*, 20739–20753. [CrossRef]
88. Riad, R.; Harba, R.; Douzi, H.; Ros, F.; Elhajji, M. Robust Fourier Watermarking for ID Images on Smart Card Plastic Supports. *Adv. Electr. Comput. Eng.* **2016**, *16*, 23–30. [CrossRef]
89. Riad, R.; Ros, F.; Gourrame, K.; Hajji, M.E.; Douzi, H.; Harba, R. A preventive and curative watermarking scheme for an industrial solution. *Multimed. Tools Appl.* **2022**, 1–29. [CrossRef]
90. Gourrame, K.; Douzi, H.; Harba, R.; Riad, R.; Ros, F.; Amar, M.; Elhajji, M. A zero-bit Fourier image watermarking for print-cam process. *Multimed. Tools Appl.* **2019**, *78*, 2621–2638. [CrossRef]
91. Prajwalasimha, S.; Chethan Suputhra, S.; Mohan, C. Performance Analysis of Combined Discrete Fourier Transformation (DFT) and Successive Division based Image Watermarking Scheme. *Int. J. Recent Technol. Eng.* **2019**, *8*, 34–39.
92. Fares, K.; Amine, K.; Salah, E. A robust blind color image watermarking based on Fourier transform domain. *Optik* **2020**, *208*, 164562. [CrossRef]
93. Chen, W.; Zhu, C.; Ren, N.; Seppänen, T.; Keskinarkaus, A. Screen-cam robust and blind watermarking for tile satellite images. *IEEE Access* **2020**, *8*, 125274–125294. [CrossRef]
94. Hsu, L.Y.; Hu, H.T. Blind watermarking for color images using EMMQ based on QDFT. *Expert Syst. Appl.* **2020**, *149*, 113225. [CrossRef]
95. Cedillo-Hernandez, M.; Cedillo-Hernandez, A.; Garcia-Ugalde, F.J. Improving dft-based image watermarking using particle swarm optimization algorithm. *Mathematics* **2021**, *9*, 1795. [CrossRef]
96. Poljicak, A.; Donevski, D.; Jelusic, P.B.; Cigula, T. Robust DFT watermarking method with gray component replacement masking. *Multimed. Tools Appl.* **2022**, *81*, 30365–30386 [CrossRef]
97. Huynh-The, T.; Hua, C.H.; Tu, N.A.; Hur, T.; Bang, J.; Kim, D.; Amin, M.B.; Kang, B.H.; Seung, H.; Lee, S. Selective bit embedding scheme for robust blind color image watermarking. *Inf. Sci.* **2018**, *426*, 1–18. [CrossRef]
98. Tan, Y.; Qin, J.; Xiang, X.; Ma, W.; Pan, W.; Xiong, N.N. A robust watermarking scheme in YCbCr color space based on channel coding. *IEEE Access* **2019**, *7*, 25026–25036. [CrossRef]
99. Zhang, L.; Wei, D. Image watermarking based on matrix decomposition and gyrator transform in invariant integer wavelet domain. *Signal Process.* **2020**, *169*, 107421. [CrossRef]
100. Al-Otum, H.M. Secure and robust host-adapted color image watermarking using inter-layered wavelet-packets. *J. Vis. Commun. Image Represent.* **2020**, *66*, 102726. [CrossRef]
101. Kumar, S.; Singh, B.K. DWT based color image watermarking using maximum entropy. *Multimed. Tools Appl.* **2021**, *80*, 15487–15510. [CrossRef]
102. Anand, A.; Singh, A.K. Joint watermarking-encryption-ECC for patient record security in wavelet domain. *IEEE Multimed.* **2020**, *27*, 66–75. [CrossRef]
103. Yuan, S.; Magayane, D.A.; Liu, X.; Zhou, X.; Lu, G.; Wang, Z.; Zhang, H.; Li, Z. A blind watermarking scheme based on computational ghost imaging in wavelet domain. *Opt. Commun.* **2021**, *482*, 126568. [CrossRef]
104. Darwish, S.M.; Al-Khafaji, L.D.S. Dual watermarking for color images: A new image copyright protection model based on the fusion of successive and segmented watermarking. *Multimed. Tools Appl.* **2020**, *79*, 6503–6530. [CrossRef]
105. Kang, X.; Chen, Y.; Zhao, F.; Lin, G. Multi-dimensional particle swarm optimization for robust blind image watermarking using intertwining logistic map and hybrid domain. *Soft Comput.* **2020**, *24*, 10561–10584. [CrossRef]
106. Swaraja, K.; Meenakshi, K.; Kora, P. An optimized blind dual medical image watermarking framework for tamper localization and content authentication in secured telemedicine. *Biomed. Signal Process. Control* **2020**, *55*, 101665.
107. Alzahrani, A.; Memon, N.A. Blind and robust watermarking scheme in hybrid domain for copyright protection of medical images. *IEEE Access* **2021**, *9*, 113714–113734. [CrossRef]
108. Zheng, P.; Zhang, Y. A robust image watermarking scheme in hybrid transform domains resisting to rotation attacks. *Multimed. Tools Appl.* **2020**, *79*, 18343–18365. [CrossRef]
109. Ernawan, F.; Ariatmanto, D.; Firdaus, A. An improved image watermarking by modifying selected DWT-DCT coefficients. *IEEE Access* **2021**, *9*, 45474–45485. [CrossRef]
110. Dargan, S.; Kumar, M.; Ayyagari, M.R.; Kumar, G. A survey of deep learning and its applications: A new paradigm to machine learning. *Arch. Comput. Methods Eng.* **2020**, *27*, 1071–1092. [CrossRef]
111. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* **2017**, *234*, 11–26. [CrossRef]
112. Naranjo-Torres, J.; Mora, M.; Hernández-García, R.; Barrientos, R.J.; Fredes, C.; Valenzuela, A. A review of convolutional neural network applied to fruit image processing. *Appl. Sci.* **2020**, *10*, 3443. [CrossRef]
113. Bagheri, M.; Mohrekesh, M.; Karimi, N.; Samavi, S.; Shirani, S.; Khadivi, P. Image watermarking with region of interest determination using deep neural networks. In Proceedings of the 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA), Virtual Event, 14–17 December 2020; pp. 1067–1072.
114. Sun, L.; Xu, J.; Liu, S.; Zhang, S.; Li, Y.; Shen, C. A robust image watermarking scheme using Arnold transform and BP neural network. *Neural Comput. Appl.* **2018**, *30*, 2425–2440. [CrossRef]
115. Cu, V.L.; Nguyen, T.; Burie, J.C.; Ogier, J.M. A robust watermarking approach for security issue of binary documents using fully convolutional networks. *Int. J. Doc. Anal. Recognit. (IJDAR)* **2020**, *23*, 219–239. [CrossRef]

116. Ingaleshwar, S.; Dharwadkar, N.V. Water chaotic fruit fly optimization-based deep convolutional neural network for image watermarking using wavelet transform. *Multimed. Tools Appl.* **2021**, 1–25. [CrossRef]
117. Sinhal, R.; Jain, D.K.; Ansari, I.A. Machine learning based blind color image watermarking scheme for copyright protection. *Pattern Recognit. Lett.* **2021**, *145*, 171–177. [CrossRef]
118. Kandi, H.; Mishra, D.; Gorthi, S.R.S. Exploring the learning capabilities of convolutional neural networks for robust image watermarking. *Comput. Secur.* **2017**, *65*, 247–268. [CrossRef]
119. Vukotić, V.; Chappelier, V.; Furon, T. Are deep neural networks good for blind image watermarking? In Proceedings of the 2018 IEEE International Workshop on Information Forensics and Security (WIFS), Hong Kong, China, 11–13 December 2018; pp. 1–7.
120. Vukotić, V.; Chappelier, V.; Furon, T. Are classification deep neural networks good for blind image watermarking? *Entropy* **2020**, *22*, 198. [CrossRef]
121. Kazemi, M.; Pourmina, M.; Mazinan, A. Analysis of watermarking framework for color image through a neural network-based approach. *Complex Intell. Syst.* **2020**, *6*, 213–220. [CrossRef]
122. Zear, A.; Singh, A.K.; Kumar, P. A proposed secure multiple watermarking technique based on DWT, DCT and SVD for application in medicine. *Multimed. Tools Appl.* **2018**, *77*, 4863–4882. [CrossRef]
123. Zear, A.; Singh, A.K.; Kumar, P. Robust watermarking technique using back propagation neural network: A security protection mechanism for social applications. *Int. J. Inf. Comput. Secur.* **2017**, *9*, 20–35. [CrossRef]
124. Huynh-The, T.; Hua, C.H.; Tu, N.A.; Kim, D.S. Robust image watermarking framework powered by convolutional encoder-decoder network. In Proceedings of the 2019 Digital Image Computing: Techniques and Applications (DICTA), Perth, Australia, 2–4 December 2019; pp. 1–7.
125. Mellimi, S.; Rajput, V.; Ansari, I.A.; Ahn, C.W. A fast and efficient image watermarking scheme based on Deep Neural Network. *Pattern Recognit. Lett.* **2021**, *151*, 222–228. [CrossRef]
126. Sinhal, R.; Ansari, I.A.; Jain, D.K. Real-time watermark reconstruction for the identification of source information based on deep neural network. *J. Real-Time Image Process.* **2020**, *17*, 2077–2095. [CrossRef]
127. Zhong, X.; Huang, P.C.; Mastorakis, S.; Shih, F.Y. An automated and robust image watermarking scheme based on deep neural networks. *IEEE Trans. Multimed.* **2020**, *23*, 1951–1961. [CrossRef]
128. Hamamoto, I.; Kawamura, M. Neural watermarking method including an attack simulator against rotation and compression attacks. *IEICE Trans. Inf. Syst.* **2020**, *103*, 33–41. [CrossRef]
129. Mun, S.M.; Nam, S.H.; Jang, H.; Kim, D.; Lee, H.K. Finding robust domain from attacks: A learning framework for blind watermarking. *Neurocomputing* **2019**, *337*, 191–202. [CrossRef]
130. Zhu, L.; Wen, X.; Mo, L.; Ma, J.; Wang, D. Robust location-secured high-definition image watermarking based on key-point detection and deep learning. *Optik* **2021**, *248*, 168194. [CrossRef]
131. Lee, J.E.; Seo, Y.H.; Kim, D.W. Convolutional neural network-based digital image watermarking adaptive to the resolution of image and watermark. *Appl. Sci.* **2020**, *10*, 6854. [CrossRef]
132. Mun, S.M.; Nam, S.H.; Jang, H.U.; Kim, D.; Lee, H.K. A robust blind watermarking using convolutional neural network. *arXiv* **2017**, arXiv:1704.03248.
133. Ahmadi, M.; Norouzi, A.; Karimi, N.; Samavi, S.; Emami, A. ReDMark: Framework for residual diffusion watermarking based on deep networks. *Expert Syst. Appl.* **2020**, *146*, 113157. [CrossRef]
134. Kim, W.H.; Kang, J.; Mun, S.M.; Hou, J.U. Convolutional neural network architecture for recovering watermark synchronization. *Sensors* **2020**, *20*, 5427. [CrossRef] [PubMed]
135. Fierro-Radilla, A.; Nakano-Miyatake, M.; Cedillo-Hernandez, M.; Cleofas-Sanchez, L.; Perez-Meana, H. A robust image zero-watermarking using convolutional neural networks. In Proceedings of the 2019 7th International Workshop on Biometrics and Forensics (IWBF), Cancun, Mexico, 2–3 May 2019; pp. 1–5.
136. Ge, S.; Xia, Z.; Tong, Y.; Weng, J.; Liu, J. A Screen-Shooting Resilient Document Image Watermarking Scheme using Deep Neural Network. *arXiv* **2022**, arXiv:2203.05198.
137. Guo, S.; Zhang, T.; Qiu, H.; Zeng, Y.; Xiang, T.; Liu, Y. The hidden vulnerability of watermarking for deep neural networks. *arXiv* **2020**, arXiv:2009.08697.
138. Lv, L. Smart Watermark to Defend against Deepfake Image Manipulation. In Proceedings of the 2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS), Chengdu, China, 23–26 April 2021; pp. 380–384.
139. Nguyen, T.T.; Nguyen, Q.V.H.; Nguyen, D.T.; Nguyen, D.T.; Huynh-The, T.; Nahavandi, S.; Nguyen, T.T.; Pham, Q.V.; Nguyen, C.M. Deep learning for deepfakes creation and detection: A survey. *Comput. Vis. Image Underst.* **2022**, *223*, 103525. [CrossRef]
140. Yu, Y.; Ni, R.; Zhao, Y. Mining generalized features for detecting ai-manipulated fake faces. *arXiv* **2020**, arXiv:2010.14129.