Journal of Information Hiding and Multimedia Signal Processing ©2025 ISSN 2073-4212
Ubiquitous International Volume **16**, Number **4**, December **2025**

A Comprehensive Review of Digital Watermarking techniques: Applications, Characteristics, Classification, and Related Aspects

Ali Nasif Jasim, Sadiq H. Abdulhussain*

Department of Computer Engineering
University of Baghdad
10071 Al-Jadriya, Baghdad, Iraq
a.jassim1805@coeng.uobaghdad.edu.iq; sadiqhabeeb@coeng.uobaghdad.edu.iq

*Corresponding author: Sadiq H. Abdulhussain

Received July 7, 2025, revised October 1, 2025, accepted October 3, 2025.

Abstract. *Watermarking has risen to be an efficient method for safeguarding digital content against unauthorized access, duplication, and modification. Over time, thorough research has been done for the purpose of enhancing its robustness, imperceptibility, and security. In this paper, a comprehensive literature review is presented to encompass the fundamental principles, characteristics, classification methods, and real-world applications of watermarking. Numerous embedding strategies, encryption techniques, and attack scenarios are examined to provide a complete understanding of the resilience and performance of watermarking systems. Beyond reviewing existing developments of watermarking, this paper explores the impact of deep learning on modern and advanced watermarking techniques which highlights recent progress in adaptive and intelligent frameworks for watermarking and outlines prospective directions for future exploration, offering valuable insights for both academic researchers and the professionals in this field. Ultimately, this paper serves as a key reference for understanding evaluation metrics and loss functions used to optimize watermarking embedding and extraction. This review merge current studies to guide practitioners and researchers in the latest digital watermarking advancements, facilitating further innovation.*
Keywords: Digital watermarking, Watermarking attacks, Transform domain, Deep learning.

1. **Introduction.** Due to the widespread digital content and the rise of digital sharing, deepfakes, and AI-generated content, the protection of intellectual property (IP) and the assurance of data integrity have gained paramount importance. Ensuring authenticity and preventing unauthorized use has become increasingly difficult, and this is where watermarking comes into place. Nowadays, watermarking has become critical for protecting ownership, verifying content integrity, and maintaining trust in digital media.

A digital watermark is an excellent way to secure personal data and prevent unauthorized distribution of multimedia files. This technique involves the insertion of hidden but detectable information into the multimedia content, which helps in authentication, copyright protection, and tamper detection. The watermarking techniques have improved greatly over the last decades; they have been significantly advanced in adding new methodologies to fight off emerging security threats and attack strategies.

This paper systematically explores watermarking, covering its applications, characteristics, classifications, and common embedding and extraction methods. It focuses on robustness through encryption and the role of deep learning in current watermarking systems. The paper also examines the impact of scaling factors, the embedding area of watermarks in images, and evaluation performance metrics that determine the success of watermarking systems. Moreover, through this literature review, knowledge gaps will be filled and researchers will be provided with a clear explanation of the current trends, challenges, and future directions in the field of digital watermarking. Knowledge obtained is crucial at all levels of watermarking processes and systems development as it will form the basis of development not only for more reliable and effective systems but also for more protected data and products in the global interconnectedness.

Figure 1 shows a classification chart that provides an organized overview of the topics covered by the research. The figure also serves as a visual aid for readers, helping them understand the general idea of the research and discover the relationships between the various theories presented within it.

FIGURE 1. A structured overview of the topics covered in this research.

2. **Related works: Strengths and Weaknesses.** Over the years, there has been a lot of work that has investigated digital watermarks from the perspectives of robustness, imperceptibility, security, and implementation details. Many of these contributions have significantly advanced the field, introducing innovative techniques with unparalleled strengths. These strengths include

high capacity, attack resistance, or computational efficiency, depending on each work's focus. While these contributions have significantly accelerated the field, they often have limitations. For example, they are limited to the use of specific techniques or lack a comprehensive analysis of multiple embedding and attack scenarios.

This review highlights strengths of existing works and identifies gaps, such as effects of encryption strategies and the balance between robustness and accuracy. The analysis highlighting these strengths and gaps is presented in Table 1.

TABLE 1. Strengths and Gaps of recent watermarking literature.

Reference	Strengths	Gaps
[1]	• Addresses ownership and copyright protection. • Maintains imperceptibility. • Uses semi-blind approach. • Tackles false positives. • Embeds principal component of watermark.	• Ignores attack types. • No discussion on capacity (BPP). • Omits scaling factor usage. • No computational analysis. • No encryption techniques used.
[2]	• Secures medical images and EPR. • Maintains visual and diagnostic quality. • Categorizes MIWT into four classes. • Reviews key evaluation metrics. • Suggests future innovations.	• Ignores attack types. • No capacity metrics (BPP). • Lacks scaling discussion. • No efficiency analysis. • No deep learning usage.
[3]	• Focuses on tamper detection and recovery. • Ensures imperceptibility. • Describes hybrid approaches. • Supports tamper localization. • Highlights optimization needs.	• Ignores attack resilience. • Lacks BPP metrics. • Omits scaling factor usage. • No timing/efficiency analysis. • No encryption discussed.
[4]	• Uses privacy-preserving techniques. • Maintains data integrity. • Aligns with watermarking goals. • Identifies security needs. • Proposes secure transmission strategies.	• Ignores attack resistance. • Lacks capacity metrics (BPP). • No mention of scaling. • No timing/efficiency data. • Omits encryption use.
[5]	• Focuses on copyright protection. • Ensures imperceptibility and robustness. • Discusses static/adaptive scaling. • Emphasizes transform domain methods. • Addresses robustness, time, and quality.	• Lacks attack details. • No BPP discussion. • Limited on scaling usage. • No PSNR/BER analysis. • Lacks encryption techniques.

3. **Watermarking Applications.** Watermarking is one of the most widely used digital rights management techniques that is applied to many applications with the aim of preventing unauthorized use, ensuring integrity, and enabling traceability. As technology is getting more and more advanced, watermarking is still an essential part of secure communication and digital rights management [6]. Some important applications of watermarking are shown in Figure 2.

In the following, we will provide a brief description of each applications and as follows:
Copyright Protection: Copyright protection in watermarking refers to embedding hidden information into digital media in order to prevent unauthorized use, duplication, or distribution

FIGURE 2. Some important applications of watermarking.

that ensures rightful ownership and legal protection [7].
Telemedicine: In telemedicine, watermarking embeds hidden data in medical images to ensure integrity, authenticity, confidentiality, and traceability [8].
Military: In military applications, watermarking embeds hidden data in digital media to ensure integrity, authenticity, and confidentiality in defense operations [6].
Remote Sensing: Watermarking in remote sensing refers to embedding hidden information into data to ensure data integrity, authenticate sources, and protect copyrights [9].
Transaction Tracking: In watermarking, transaction tracking embeds hidden data to monitor content distribution, verify authenticity, and control IP [10].
Tamper Detection: In watermarking, tamper detection embeds hidden information in digital media to ensure content integrity and authenticity, enabling identification and localization of unauthorized alterations [11].
Image Authentication: In watermarking, image authentication embeds hidden information into an image to verify originality and detect unauthorized alterations, ensuring integrity and identifying tampered regions [12].
Digital Forensics: Digital Forensics in watermarking involves the use of digital watermarking techniques to verify the authenticity, integrity, and origin of digital media. By embedding unique identifiers or authentication data into media files, investigators can detect tampering, trace the source of unauthorized distributions, and establish the provenance of digital content [13].
Content Privacy: Content Privacy in watermarking refers to embedding information within digital media to protect sensitive data, ensuring that unauthorized access or distribution is prevented. This technique not only protects the content but also maintains the confidentiality of the information which prevents misuse or unauthorized sharing [14].
Remote Education: In remote education, digital watermarking is employed to embed hidden information within digital educational resources such as videos, and documents to ensure content integrity, authenticate users, and protect IP [15].
E-Voting: In electronic voting, watermarking is employed to embed hidden information within

electronic ballots or voter data to enhance security, ensure authenticity, and maintain the integrity of the voting process. This technique helps in preventing fraud, verifying voter identities, and ensuring that votes are accurately recorded and counted [16].
Health Care: In the healthcare sector, watermarking involves embedding hidden information within medical data, such as images or electronic health records, to ensure their authenticity, integrity, and confidentiality. This technique protects patient information, verifies data origin, and detects unauthorized alterations. Therefore, it maintains trust in medical communications and diagnoses [17].

4. **Watermarking Characteristics.** The term 'Watermarking' is described and defined according to special features such as robustness, invisibility, and privacy. A strong and resistant watermark can withstand attacks while remaining invisible to preserve the quality of the content. Its efficiency depends on balancing durability and distortions, along with reliable and verifiable embedding. Some characteristics to consider when evaluating watermarking efficiency are shown in Figure 3.

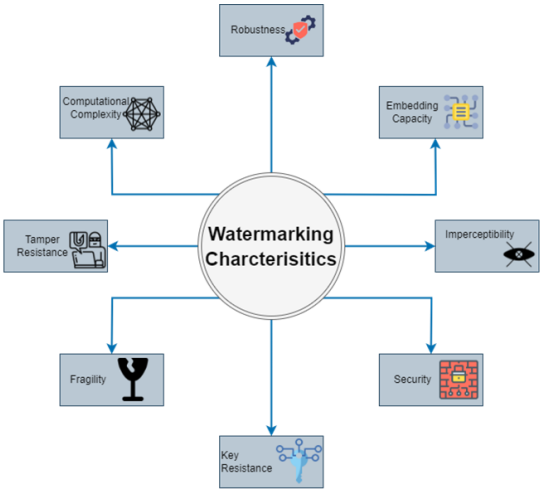

FIGURE 3. Various important characteristics of an effective watermark.

5. **Classification of Watermarking Methods.** Watermarking methods are classified into three main categories based on accessibility, embedding domain, and key attributes. Accessibility-based watermarking includes public schemes, which are openly available, and private schemes, restricted to authorized users. Embedding domain classification differentiates spatial domain techniques that modify pixel values directly, transform domain methods embedding watermarks in frequency components, and hybrid approaches combining both. Attribute-based classification considers security (robust vs. fragile), extraction (blind vs. non-blind), and visibility (visible vs. invisible). These classifications help determine a watermarking method's suitability for different applications. Figure 4 illustrates the three basic categories of watermarking methods and their sub-classification methods.

5.1. **Based on Accessibility.** One of the categories that assists in classifying the watermarking system. It enables the authorized users to access the embedded watermark. It can be further classified into:

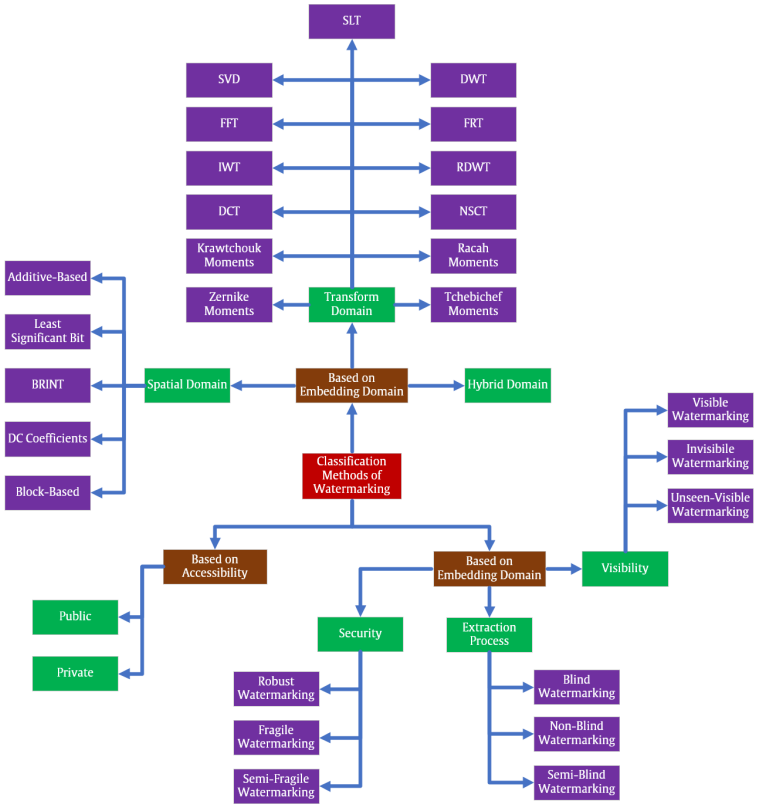

FIGURE 4. The three basic categories of watermarking methods and their sub-classification methods.

5.1.1. *Public.* In this scheme, the watermark is accessible to everyone or a specific group, making it easily modifiable and less secure [18].

5.1.2. *Private.* In this scheme, the watermark is accessible to only verified or authentic users. This watermark is embedded using a secret key which is shared only among authenticated individuals.

5.2. **Based on Embedding Domain.** The embedding domain is another category under which the watermarking techniques can be classified. This domain affects the robustness, imperceptibility, and effectiveness of the watermark. Choosing an embedded domain determines the resistance of the watermark to attack/manipulation and the quality of the original content. Knowing the various embedding domains helps select the right watermarking technique for an application [19].

5.2.1. *Spatial Domain.* In digital watermarking, the spatial domain directly manipulates the pixel values of the image to embed a watermark. Such a method changes the intensity or color of specific pixels to reflect the desired information. Spatial domain techniques are simple and computationally fast techniques that are very common in many applications. They are, however, less robust against some image processing operations and attacks than the transform domain methods [20]. They are divided into:

- In digital image watermarking, the Least Significant Bit (LSB) method embeds watermark data by altering only the least significant bits of pixel values, which minimally affects visual quality and is practically undetectable. While LSB is computationally efficient and preserves high image fidelity, it is generally less robust against image processing attacks like compression or noise addition [21, 22].
- Block-Based: In the block-based method, the process involves partitioning the image into non-overlapping blocks and placing the watermark in each segment. This allows localized embedding, which may improve robustness against localized attacks and speed up processing [23].
- Additive-Based: In digital watermarking, the additive-based spatial domain method embeds a watermark as pseudo-random noise by directly adding it to pixel values. Its main advantage is simplicity, but it may be less resilient to attacks like compression or noise addition compared to other methods [24].
- DC Coefficients: In digital image watermarking, the DC coefficients method in the spatial domain involves embedding watermark information into the Direct Current (DC) components of an image. The DC component represents the average color or intensity of a block of pixels, making it less sensitive to minor alterations and compression, thereby enhancing the robustness of the watermark [25, 26].
- BRINT: The BRINT method, belonging to the spatial domain, enhances texture feature extraction by improving rotation invariance and noise robustness. Its "averaging before binarization" strategy reduces noise sensitivity, making it ideal for watermarking where robustness against distortions is crucial. BRINT ensures stable feature representation under transformations, improving watermark embedding and extraction reliability [27].

5.2.2. *Transform Domain.* Another domain in digital image watermarking is the transform domain, where an image is transformed from its spatial domain to the frequency domain for embedding watermark information. This approach basically exploits the mathematical characteristics of transforms such as the Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), or Discrete Wavelet Transform (DWT) to embed watermarks that are less sensitive to the human visual system (HVS) and more robust against usual image processing (IP) operations [28, 29]. The transform domain can be divided into:

- Discrete Cosine Transform (DCT): DCT is an image transformation technique from the spatial domain to the frequency domain for embedding watermark information into frequency components in digital image watermarking. With this approach, the properties of DCT are exploited to balance imperceptibility with robustness. Using DCT, spatial data can be transformed into the frequency domain and reverted back to the spatial domain through the Inverse Discrete Cosine Transform (IDCT) [30].
- DWT (Discrete Wavelet Transform): The Discrete Wavelet Transform (DWT) is a transform domain technique that decomposes an image into four frequency sub-bands—low-low (LL), low-high (LH), high-low (HL), and high-high (HH)—facilitating watermark embedding that balances imperceptibility and robustness. The LL sub-band holds the image's approximation (primary structure), while LH, HL, and HH contain details like edges and textures. This decomposition can be applied recursively to LL, forming a multi-level hierarchy. DWT coefficients in the LL sub-band typically have higher magnitudes than in other sub-bands [31–33].
- SVD (Singular Value Decomposition): Singular Value Decomposition (SVD) is a fundamental matrix factorization technique in linear algebra, widely utilized in digital image watermarking within the transform domain like compression, information hiding and noise reduction [34].

- SLT (Slantlet Transform): The Slantlet Transform (SLT) is a linear transform that extends the discrete wavelet transform (DWT) by providing improved time localization and two zero moments, making it particularly advantageous for applications like image watermarking [35].
- FFT (Fast Fourier Transform): The Fast Fourier Transform (FFT) is an algorithm for computing the DFT (and its inverse). Direct computation of DFT requires time, whereas FFT reduces this time and is much faster for large datasets. This is done efficiently by factoring the DFT into smaller DFTs (see Cooley-Tukey algorithm). The watermark must be embedded into a host signal, such as audio or video, in digital watermarking. In such systems, the frequency components of the host signal are manipulated by FFT so that the watermark is imperceptible but robust to attack [36].
- IWT (Integer Wavelet Transform): IWT is a fully discrete wavelet transform defined in the integer domain, without using floating-point precision. This property makes IWT very useful in digital watermarking systems where embedding and extracting watermarks are required very accurately [37].
- FRT (Fractional Fourier Transform): The Fractional Fourier Transform (FRT) is a generalization of the traditional Fourier Transform (FT) for signals in intermediate domains between frequency and time. Introduced by a fractionally ordered parameter α, the FRT gives a continuum of transformations corresponding to each rotation in the time-frequency plane. This property makes the FRT particularly suitable for signal processing applications like digital watermarking [38].
- RDWT (Redundant Discrete Wavelet Transform): Is an extension of traditional Discrete Wavelet Transform (DWT), which overcomes the shift variance problem of DWT. Instead of down-sampling the signal, which may lose the information as in DWT due to its shift-variant characteristic, RDWT keeps the same size as the original signal. Such redundancy enables a finer localization of the frequency components, which makes RDWT very suitable for digital watermarking applications. RDWT decomposes an image into sub-bands without down-sampling, preserving the original resolution [39, 40].
- NSCT (Non-Subsampled Contourlet Transform): NSCT is an advanced image transformation technique that provides better representation of image features than traditional wavelet transforms. It has two major components:
  - Non-Subsampled Pyramid (NSP): Decomposes an image into a low-pass sub-band and several high-pass sub-bands without down-sampling, which ensures shift invariance.
  - Non-Subsampled Directional Filter Bank (NSDFB): Directional filtering is applied to the high-pass sub bands to capture geometric structures in multiple orientations.

  NSCT is utilized in watermarking thanks to its high directional selectivity, improved geometric attack resistance, and the frequency representation that is better than that of a traditional wavelet-based method [41, 42].
- Zernike Moments: Zernike moments, introduced by Teague, are orthogonal moments derived from Zernike polynomials defined over the unit circle in polar coordinates. Widely used in image processing and pattern recognition, they provide a compact, rotation-invariant image representation. In watermarking, Zernike moments enable robust embedding against geometric transformations like rotation, scaling, and flipping [43, 44].
- Racah Moments: Racah moments are orthogonal moments derived from Racah polynomials, originally introduced by Wilson and later adapted by Zhu for discrete settings. They are valuable in image processing, pattern recognition, and signal representation due to their robust feature extraction. Their key advantage is discrete orthogonality, making them well-suited for digital image analysis and classification [45, 46].

- Tchebichef Moments: Tchebichef moments are orthogonal moments derived from discrete Tchebichef polynomials and are defined directly in the image coordinate space. Unlike traditional moments like Zernike, they require no numerical approximations or coordinate transformations, making them computationally efficient. Well-suited for object classification and image reconstruction. Tchebichef moments maintain discrete-domain orthogonality, as the underlying Tchebichef polynomials are generated using a recurrence relation [47]. This ensures minimal redundancy and improved numerical stability [48, 49].
- Krawtchouk Moments: Krawtchouk moments are discrete orthogonal moments derived from Krawtchouk polynomials linked to the binomial distribution. Unlike moments such as Zernike or Legendre, they capture local image features by varying a probability parameter, enabling region-of-interest analysis. Computationally efficient without requiring spatial normalization or numerical approximation, they maintain discrete orthogonality, reducing redundancy. Applications include image reconstruction, pattern recognition, and object classification [50].

5.2.3. *Hybrid Domain.* Hybrid domain in watermarking combines domains such as spatial, frequency, and transform domains in order to embed watermarks into digital content. This exploits the strengths of the respective domain to provide robustness and imperceptibility, which the single-domain approach cannot achieve. Hybrid watermarking techniques combine these domains to leverage the benefits of each for watermarks that are as robust as well as imperceptible [23]. Watermarking using hybrid domains combines domains such as spatial, frequency, and transform domains to embed watermarks into digital content. This exploits the strengths of the respective domain to provide robustness and imperceptibility, which the single-domain approach cannot achieve. Hybrid watermarking techniques combine these domains to leverage the benefits of each for watermarks that are as robust as well as imperceptible [51, 52]. Table 2 shows a comparison between the three domains of watermarking.

TABLE 2. Cross-Domain Comparison of Watermarking Techniques.

Domain	**Complexity**	**Robustness**	**Capacity**	**Actual Logic**
Spatial	Low	Low	Medium	Directly modifies pixel values
Transform	Medium	High	Medium	Converts data to frequency domain (e.g., DCT, DFT)
Hybrid	High	High	High	Combines spatial and frequency domain techniques

5.3. **Based on Attributes.** Security, extraction process, and visibility are three key attributes defining watermarking techniques. Security classification distinguishes methods based on resilience to attacks like removal or tampering. Extraction-based classification depends on how the watermark is retrieved—whether the original image or watermark is needed for detection. Visibility classifies the watermark as perceptible or imperceptible. These classifications help select suitable watermarking methods by balancing robustness, usability, and security.

5.3.1. *Security.* In digital watermarking, security-based classification considers the watermark's resistance to unauthorized detection, removal, or alteration. It assesses how well the watermark withstands intentional attacks on its integrity or concealment. Based on this attribute, watermarking can be categorized into [53]:

- Robust Watermarking: is a technique in which the watermark is created to resist numerous kinds of attacks or modifications, maintaining its integrity and detectability even after substantial alterations to the digital content [54].

- Fragile Watermarking: is a technique in which the watermark becomes damaged or even altered if the digital content is altered; therefore, it may be used to identify tampering and unauthorized modifications [55].
- Semi-Fragile Watermarking: is a digital watermarking technique designed to balance robustness and sensitivity, enabling it to withstand minor modifications while detecting and signaling malicious alterations [56].

5.3.2. *Extraction Process.* In digital watermarking, extraction is the process to retrieve the embedded watermark from watermarked content for the purpose of authenticity/ownership verification. This is basically the reverse of embedding and depends on the watermark technique applied. The three main types are:

- Blind Watermarking: Blind means that in the watermark extraction process, the original, unwatermarked content does not have to be accessible to retrieve the embedded watermark [57].
- Non-Blind Watermarking: is an extraction process requires access to the original unwatermarked content to retrieve the embedded watermark [58].
- Semi-Blind Watermarking: Falls between blind and non-blind methods as some partial information about the original content/embedding process is exploited and needed at the extraction [59].

5.3.3. *Visibility.* Visibility means whether the watermark is easily detectable in the digital content or whether it is obstructive to the human eye. It is subdivided into:

- Visible Watermarking: Here, the watermark is embedded in the digital content in such a way that it becomes visible to the human eye; it is normally used for branding or copyright purposes [60].
- Invisible Watermarking: The watermark is invisibly embedded in digital content to provide copyright protection and data tracking without altering its appearance [60].
- Unseen-Visible Watermarking: It is a type of digital watermarking where the watermark is invisible to the unaided eye but visible under certain conditions, combining benefits of visible and invisible watermarking while maintaining imperceptibility and robustness [61].

6. **Watermarking Optimization Methods.** Optimization methods are computational strategies designed for the purpose of identifying the most efficient solution to a given problem by either maximizing or minimizing an objective function. In digital watermarking, these techniques are essential for maintaining a balance between crucial factors such as robustness, imperceptibility, security, and capacity. By carefully fine-tuning parameters like the scaling factor, embedding location, and feature selection, optimization algorithms improve the overall effectiveness of watermarking schemes, making them more resistant to attacks and distortions [62]. Optimization methods can be broadly classified as shown in Figure 5.

6.1. **Evolutionary Based Algorithms.** Algorithms which mimic the process of evolution in nature are known as evolutionary-based algorithms. Genetic algorithms (GAs) have been widely used in watermarking with fitness functions to improve robustness and imperceptibility. Watermarking parameters are encoded as chromosomes in these algorithms which are subjected to crossover and mutation. This process continues until convergence is achieved. Researchers have combined GA with discrete cosine and wavelet transforms to optimize embedding parameters, enhancing both security and resistance to attacks. Other types of population-based metaheuristic methods, such as memetic algorithms and adaptive dimensional search, have also been used in watermarking for fine-tuning performance [62]. Examples of Evolutionary Based Algorithms include:

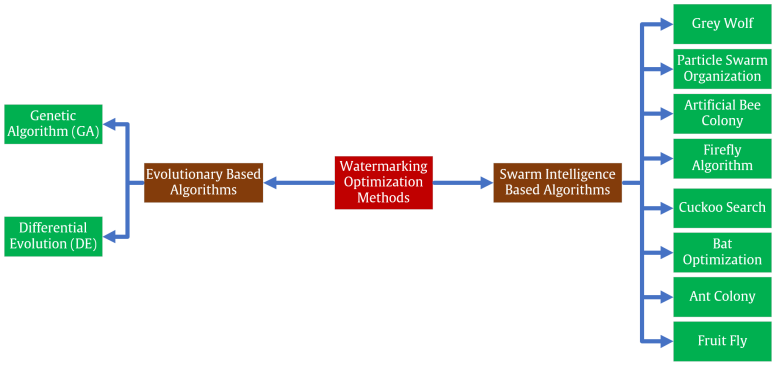

FIGURE 5. Watermarking optimization methods.

6.1.1. *Genetic Algorithm (GA).* GAs are optimization techniques based on evolution and natural selection. They have been widely used in digital watermarking to improve robustness, invisibility and security. GA functions by utilizing important evolutionary processes such as reproduction, crossover, and mutation. Watermarking involves the use of GA for optimizing embedding parameters such that the watermark is resistant to attacks and quality of image is also maintained. The procedure initiates by establishing the watermark embedding parameters, followed by the formation of an initial population of solutions (chromosomes). Each solution is evaluated using a fitness function to measure its robustness and imperceptibility. The algorithm uses crossover to share genetic information between selected solutions and mutation to avoid local optima. This process will keep repeating until convergence is reached, yielding an optimized scheme [62].

6.1.2. *Differential Evolution (DE).* Differential Evolution is an optimization algorithm that is evolution-based and belongs to the family of evolutionary algorithms. Storn and Price introduced it in 1996 and ever since then, it has grown to be widely used for solving complex optimization problems. DE algorithm functions through iterative movements, beginning with a randomly created population of candidate solutions. It evolves and advances this population through using 3 primary operations which are 'Mutation', 'Crossover', and 'Selection'. Mutation involves creating a new solution by combining existing solutions using a weighted difference strategy. As for Crossover, it allows for information to be exchanged between solutions to increase diversity. And finally, Selection ensures that the fittest individuals are carried forward to the next generation. DE is efficient for continuous optimization and has been applied in engineering, machine learning, and control systems. It offers simplicity, robustness, and can handle non-differentiable, nonlinear, and multi-modal objective functions [63].

6.2. **Swarm Intelligence Based Algorithms.** Another powerful approach to optimization is offered by swarm intelligence algorithms, inspired by collective behavior in nature. For example, ant colony optimization (ACO) algorithms, which mimic how ants converge in large numbers to find the shortest path to a food source, also metaphorically apply this kind of logic to detect optimal positions for watermark embedding. Inspired by the notion that individual birds in a flock or fish in a school adjust their positions based on personal and global best solutions, particle swarm optimization (PSO) effectively leads convergence to an optimal embedding strategy. Other methods include bee algorithms, cuckoo search, and firefly algorithms. All these methodologies belong to the class of swarm algorithms that maintain image fidelity while enhancing watermark robustness. Some of these algorithms, when combined with wavelet transforms, have been found

to withstand attacks on images, even from advanced watermark attackers [62]. Examples of swarm intelligence-based algorithms include the following:

6.2.1. *Grey Wolf.* Grey Wolf is a swarm intelligence-based optimization technique inspired by the social hierarchy and hunting behavior of grey wolves. Introduced by Mirjalili et al., GWO models the leadership structure of a wolf pack, where the alpha, beta, and delta wolves guide the search process, while omega wolves follow their lead. The algorithm mimics the pack's hunting strategy, which involves tracking, encircling, and attacking prey. During optimization, wolves adjust their positions relative to the best solutions found so far, balancing exploration and exploitation. The encircling mechanism helps wolves converge on promising areas, while adaptive coefficient vectors enable dynamic adjustments to enhance solution diversity. GWO has been successfully applied to various optimization problems, showing competitive performance compared to algorithms like PSO and GAs [64].

6.2.2. *Particle Swarm Optimization.* PSO is a swarm intelligence optimization method based on the phenomenon of collective bird flocking or fish schooling behavior. Conceptualized by Kennedy and Eberhart, PSO belongs to the family of stochastic optimization algorithms. A swarm of particles, wherein each particle represents a candidate solution, moves in the solution space. Movement is updated based on a particle's own best position (experience) and the global best position found by any particle in the swarm (best neighborhood experience). The movement of a particle is governed primarily by updating its velocity. This cognitive and social inspiration plus somewhat carefully chosen element of stochastic choice keep it exploring. Over time, after many iterations, the swarm will eventually "feel" its way to a good solution. Some of the main successes found in various fields are achieved in function optimization, machine learning, control systems, and engineering design; PSO is simple, fast, and it is easy to implement as it is also capable of handling nonlinear and multi-modal problems [65].

6.2.3. *Artificial Bee Colony.* The Artificial Bee Colony (ABC) algorithm was developed by Karaboga in 2005 as a swarm-intelligence-based optimization algorithm that mimics the well-known foraging behavior of real honeybees. The ABC algorithm uses three types of bees: employed bees, onlooker bees, and scout bees. Employed bees exploit food sources and share findings with onlooker bees via waggle dance, enabling onlookers to probabilistically select sources based on nectar amount. Meanwhile, scout bees search for new food sources in the neighborhood of abandoned sources with a certain probability. In ABC, food source positions represent potential solutions. Their quality is nectar quality, a measure of fitness turned into fitness landscapes for solving minimization problems. Exploration and exploitation phases gradually refine solutions, making the algorithm efficient for continuous, combinatorial, and multi-objective optimization problems. ABC is widely used across domains for its simplicity and efficiency, including engineering, image processing, data mining, and neural network training [66].

6.2.4. *Fire Fly Algorithm.* A swarm-intelligence optimization technique invented by Xin-She Yang in 2008, inspired by the flashing behavior of fireflies used for mating communication and predation avoidance. Each firefly represents a potential solution, with its brightness indicating the solution's quality or fitness. The basic principle of FA is that less bright fireflies are attracted to brighter ones. The degree of such attractiveness diminishes with the increase in the distance between fireflies. Such an attraction mechanism efficiently guides the swarm to explore the search space while converging toward the optimal solution. FA combines exploration (global search) and exploitation (local search), balancing them through parameters like the absorption coefficient and randomness. Due to its ease of implementation and effectiveness, FA has been applied successfully to a wide variety of optimization problems. These range from optimal design to image processing, from tuning of neural networks to problems of combinatorial optimization [67].

6.2.5. *Cuckoo Search.* The Cuckoo Search algorithm was developed by Xin-She Yang and Suash Deb in 2009 as a heuristic optimization algorithm based on the obligate brood parasitism behavior of some cuckoo species. These birds lay their eggs in the nests of other host birds, sometimes removing the host's eggs to maximize their chances of survival. In the Cuckoo Search algorithm, the "cuckoo egg" represents a candidate solution. Each generation selects the best solutions for offspring nests, while poor solutions are abandoned and replaced with new nests to maintain population size. It tries to use both Lévy flight and local random walk for the global search. Lévy flights might be useful for increasing the exploration of the search space with large jumps. In comparison to other optimization techniques like GAs and PSO, this technique has demonstrated greater efficiency in finding global optima due to its balanced exploitation and exploration mechanisms. It has been successfully applied to engineering design, machine learning, structural optimization, and other complex optimization problems [68].

6.2.6. *Bat Optimization.* The Bat Optimization Algorithm is a swarm intelligence-based optimization technique that was proposed by Xin-She Yang and is based on the echolocation behavior of bats. Bats use sound pulses to navigate and detect prey, adjusting their frequency, loudness, and pulse rate dynamically. In this technique, each bat represents a potential solution in the search space, which moves towards optimal solutions by updating both its velocity and position based on the best global and local solutions found. This algorithm basically integrates both exploration (which is global search) and exploitation (which is local search) mechanisms. This in turn balances randomness and convergence by adapting parameters such as loudness and pulse rate over iterations. This algorithm has been successfully applied to various engineering optimization problems, outperforming many traditional optimization techniques due to its adaptability and efficiency [69].

6.2.7. *Ant Colony.* Ant Colony Optimization (ACO) stands as one of the leading swarm intelligence methods that have been used to solve combinatorial optimization problems. ACO is inspired by real ants' foraging behavior, where pheromone trails help communicate and find the shortest paths to food. Algorithms mimicking ACO use probability to construct solutions, similar to how artificial ants traverse paths to food sources. In this process, pheromone levels are updated based upon the quality of solutions constructed by the ants, thereby reinforcing good paths. Hence, it fosters an adaptive search process through the slow gain of strength in the more promising solutions. Over time, a natural balance of explorative behavior and exploitative behavior is achieved by a straightforward evaporation of pheromones to avoid stagnation but an intensification of the best solutions. This algorithm has been successfully applied to problems like the Traveling Salesman problem, scheduling, and network routing. Demonstrating its effectiveness in dynamic and complex optimization scenarios [70].

6.2.8. *Fruit Fly.* The Fruit Fly Optimization Algorithm (FOA) is a swarm intelligence-based optimization technique that simulates the foraging behavior of fruit flies (Drosophila). Fruit flies have a keen sense of smell and vision, allowing them to locate food sources efficiently. In this algorithm, individuals in the population represent potential solutions, and these solutions update their positions in iterative manner in the solution space based on smell-based and vision-based searches. This algorithm consists of two main phases which are the smell phase, where solutions are generated based on heuristic search strategies, and the vision phase, where the best solution found so far is used to guide the next search direction. There are variations of FOA, such as the Q-learning-enhanced FOA (QFOA), which incorporate reinforcement learning techniques to dynamically select high-quality neighborhood structures, improving the algorithm's exploration and exploitation capabilities. FOA has been successfully applied in various fields, including scheduling, power load forecasting, and optimization problems [71].

7. **Watermarking Attacks.** Digital watermarking embeds information into digital media to assert ownership, track usage, and ensure content integrity. However, these watermarks can be vulnerable to attacks that aim to remove or alter the embedded information. Understanding these attacks is crucial to develop robust watermarking systems. Figure 6 shows classification of watermarking attacks.

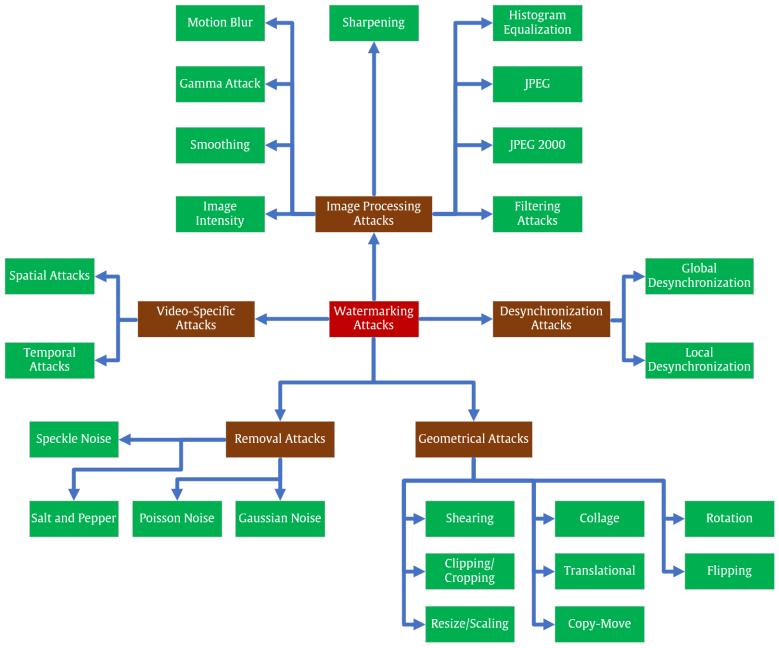

FIGURE 6. Classification of attacks on watermarking.

7.1. **Image Processing Attacks.** Image processing attacks involve modifications like filtering, resizing, or color adjustments applied to watermarked images to remove, degrade, or alter the embedded watermark. They can be divided into:

7.1.1. *Image Intensity.* Image intensity attacks involve altering the brightness or contrast of an image to disrupt or remove the embedded watermark. These attacks can be intentional, aiming to breach copyright or authenticity, or unintentional, caused by standard image processing operations [72]. Figure 7b shows an intensity attack on an image.

7.1.2. *Smoothing.* A smoothing attack uses image processing techniques like blurring or averaging to reduce an image's high-frequency components. This process can diminish or eliminate the embedded watermark, especially if it's situated in these high-frequency areas [73]. Figure 7c shows an image smoothing attack.

7.1.3. *Gamma Attack.* A gamma attack is a type of attack that uses gamma correction on a watermarked image to remove or ruin the embedded watermark. Gamma correction is the nonlinear operation that changes image luminance (brightness and contrast). Changes in the gamma value affect how the image looks visually, which makes the watermark less detectable [74]. Figure 7d shows a gamma attack.

7.1.4. *Motion Blur.* Another attack type applies a motion blur filter to a watermarked image, simulating camera movement during capture. This geometric transformation can distort or completely remove the embedded watermark, making it less detectable [75]. Figure 7e shows a motion blur attack on an image.

7.1.5. *Sharpening.* A sharpening attack applies a filter to enhance edges and details in a watermarked image. Sharpening improves visual clarity but also distorts or remove the embedded watermark [76]. Figure 7f depicts a sharpening attack.

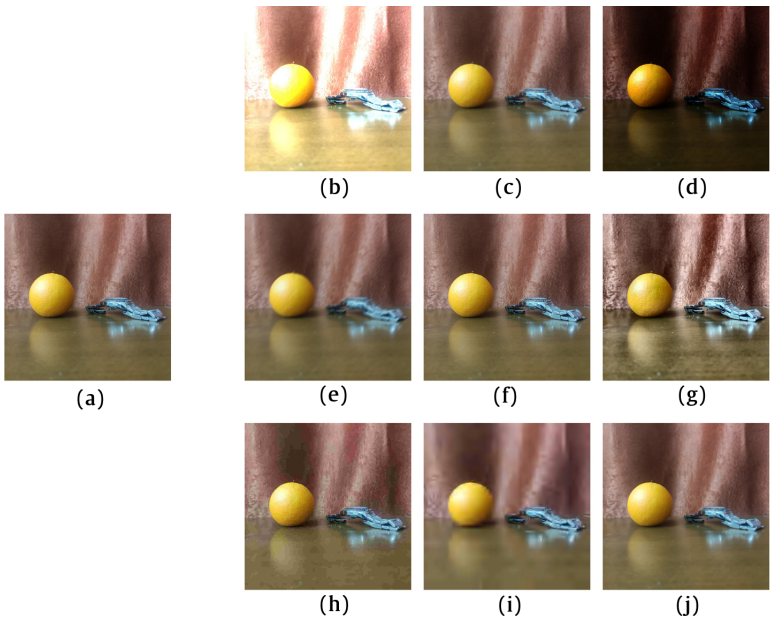

FIGURE 7. Image processing attacks, (a) Original image, (b) Image intensity, (c) Smoothing, (d) Gamma, (e) Motion blur, (f) Sharpening, (g) Histogram equalization, (h) JPEG compression, (i) JPEG 2000 compression, and (j) Filtering.

7.1.6. *Histogram Equalization.* When histogram equalization applied to watermarked image, it increases the contrast of an image by distributing its intensity levels more evenly, which may skew or remove the embedded watermark [77]. Figure 7g shows histogram equalization attack.

7.1.7. *JPEG Compression.* JPEG is a lossy image compression algorithm that exploits coding redundancy, inter-pixel redundancy, and psycho-visual redundancy. It applies the DCT to 8×8 pixel blocks, followed by quantization, which suppresses higher frequency components [78]. Figure 7h shows JPEG Compression attack.

7.1.8. *JPEG 2000 Compression.* JPEG 2000, based on wavelet transform, was developed to overcome JPEG's limitations. Instead of dividing the image into blocks, it applies wavelet-based encoding, which avoids blocking artifacts [78]. Figure 7i shows JPEG 2000 Compression attack on an image.

7.1.9. *Filtering Attacks.* Filtering attacks apply various filters to a watermarked image to degrade or remove the embedded watermark. Filters like averaging, median, and Gaussian are commonly used for noise reduction and image enhancement. However, when applied to watermarked images, they unintentionally or maliciously impair the watermark's integrity [77]. Figure 7j shows Filtering attacks.

7.2. **Removal Attacks.** Removal attacks aim to completely eliminate the embedded watermark from an image without degrading the host image's quality. These attacks are critical because they directly threaten watermarking's core goals of protecting IP and verifying authenticity [79]. It is divided into:

7.2.1. *Salt and Pepper Noise.* Salt and pepper noise introduces random white and black pixels into an image, resembling grains of salt and pepper. Noise like this can make the watermark less detectable and affect its integrity. In one study, salt and pepper noise attacks effectively removed a watermark embedded in a digital image [80]. Figure 8b shows a salt and pepper attack on an image.

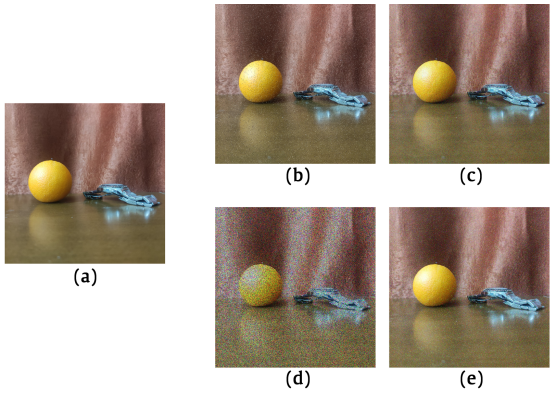

FIGURE 8. Removal attacks: (a) Original image. (b) Salt and pepper noise, (c) Gaussian noise, (d) Speckle noise, and (e) Poisson noise.

7.2.2. *Gaussian Noise.* Gaussian noise adds random pixel variations following a Gaussian distribution, which can obscure the watermark and hinder its extraction. Research indicates that Gaussian noise attacks can degrade watermark quality depending on noise level [81]. Figure 8c shows a Gaussian Noise attack.

7.2.3. *Speckle Noise.* Speckle noise is a granular noise produced by interference patterns. This causes graininess in images and may obscure the watermark. Studies show that speckle noise attacks affect watermark detection when noise is high [82]. Figure 8d shows a Speckle noise attack.

7.2.4. *Poisson Noise.* Poisson noise arises from photon count variations during image capture, causing pixel intensity fluctuations that can obscure the watermark. Research indicates that Poisson noise attacks can affect watermark robustness depending on noise characteristics [83]. Figure 8e shows a Poisson Noise attack.

7.3. **Geometric Attacks.** Geometric attacks are spatial changes like rotation, scaling, translation, and cropping. These attacks disrupt the synchronization of the embedded watermark, making it undetectable and ineffective [84]. Examples of geometric attacks are:

7.3.1. *Resize/Scaling Attack.* Resizing (scaling) changes the image dimensions and alters the watermark relative to the image size. This skews or blurs the watermark and makes it less visible. For example, scaling an image 2 times reduces the watermark by half, making it undetectable. Figure 9b shows Resizing attack.

7.3.2. *Rotation Attack.* Rotation involves turning an image around and changing its orientation. It causes the watermark to be off-center with the host image. This misalignment may prevents a detection/extraction of the watermark. Figure 9c shows Rotation attack.

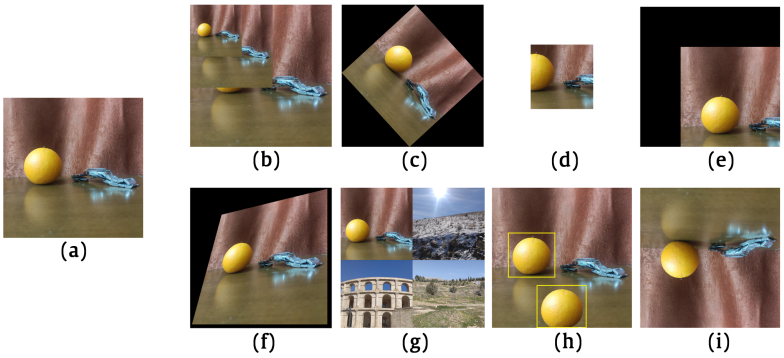

FIGURE 9. Geometric attacks. (a) Original image. (b) Resize/scaling attack. (c) Rotation attack. (d) Clipping/cropping attack. (e) Translation attack. (f) Shearing attack. (g) Collage attack. (h) Copy-move attack. (i) Flipping attack.

7.3.3. *Clipping/Cropping Attack.* Clipping removes parts of the image, such as cropping out a watermark region. This may cause partial or complete loss of the watermark if the clipped area contains critical watermark information. Figure 9d shows Clipping/Cropping attack on an image.

7.3.4. *Translation Attack.* Translation shifts the image in space and moves the watermark to a different location. This displacement causes synchronization errors, making the watermark difficult to detect or extract. Figure 9e shows Translation attack.

7.3.5. *Shearing Attack.* Shearing slants the image, which can change the look of the watermark. Such a transformation may make the watermark less recognizable or unrecognizable. Figure 9f shows Shearing attack.

7.3.6. *Collage Attack.* In collage attacks, the watermark gets embedded into several images to create a new image from multiple sources. It confuses detection algorithms and breaks the watermark integrity. Figure 9g shows Collage attack.

7.3.7. *Copy-Move Attack.* Copy-move attacks involve pasting a region from the same image onto another location. This results in false positives for watermark detection as the duplicated region contains parts of the watermark. Figure 9h shows Copy-Move attack.

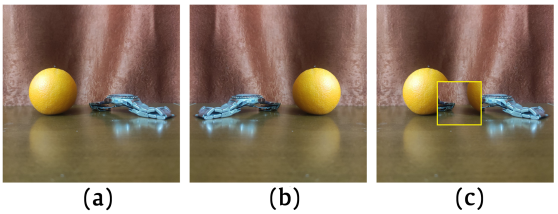

FIGURE 10. Desynchronization attacks: (a) Original image, (b) Global desynchronization, and (c) Local desynchronization.

7.3.8. *Flipping Attack.* Flipping mirrors horizontally or vertically changes the orientation of the watermark. Such a transformation causes the watermark to not match detection algorithms. Which makes it difficult to extract. Figure 9i shows Flipping attack.

7.4. **Desynchronization Attacks.** Desynchronization attacks disrupt watermark detection by modifying the spatial or temporal alignment between the embedded watermark and the host content so that the watermark is not easily detected or extracted [85]. They can be divided into:

7.4.1. *Global Desynchronization.* Global Desynchronization attack is a type of watermarking attack where the entire image or video is shifted or altered in a way that disrupts the alignment of the watermark, making it difficult to detect or extract the watermark accurately. Figure 10b shows Global Desynchronization attack.

7.4.2. *Local Desynchronization.* Local Desynchronization attack is a type of watermarking attack where only specific regions of the image or video are altered or shifted, disrupting the alignment of the watermark in those localized areas. Figure 10c shows Local Desynchronization attack on an image.

7.5. **Video Specific Attacks.** Video watermarking is more complex than image watermarking due to the additional temporal dimension. Video-specific attacks exploit this dimension by manipulating frames, motion, or encoding parameters to degrade or remove the watermark [60]. These attacks can be classified into:

7.5.1. *Spatial Attacks.* Spatial attacks are a type of video-specific attack where modifications are made to the spatial domain of individual video frames, such as cropping, resizing, or altering regions, which can disrupt or remove the embedded watermark. Figure 11a shows a spatial attack on a video.

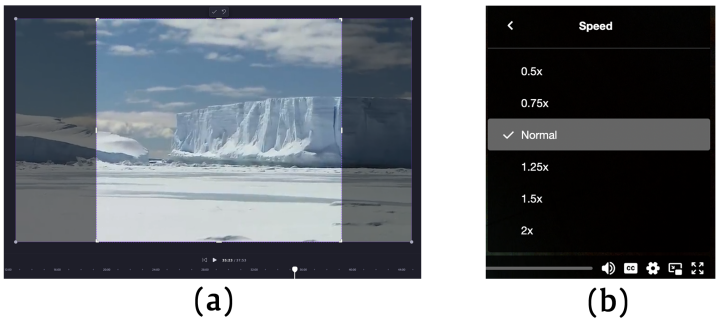

FIGURE 11. Video specific attacks: (a) Spatial, and (b) Temporal.

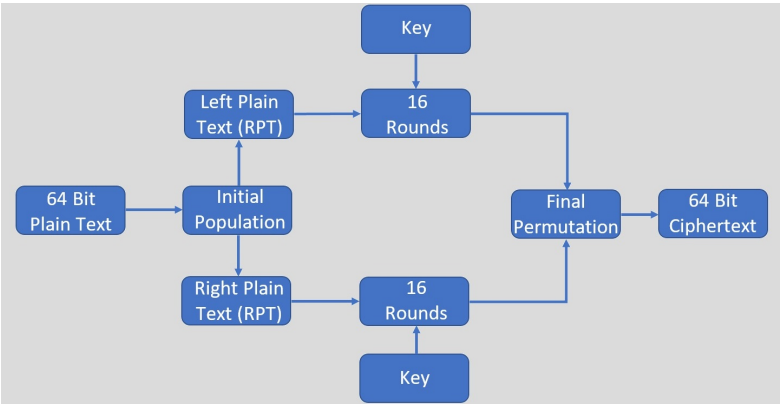

FIGURE 12. DES algorithm.

7.5.2. *Temporal Attacks.* Temporal attacks are video-specific modifications to the temporal domain, such as altering frame sequence, dropping frames, or changing playback speed, which can disrupt or remove the embedded watermark. Figure 11b shows a temporal attack on a video.

8. **Watermarking Scaling Factors.** Scaling factors are parameters that adjust the watermarking system to size or resolution changes, ensuring the watermark remains effective and detectable [60, 86].

1. Static Scaling Factors: are fixed parameters used in watermarking systems to handle changes in content size or resolution, ensuring consistent watermarking performance without adjustment during processing .
2. Adaptive Scaling Factors: are dynamic parameters in watermarking that adjust in response to changes in content size or resolution, optimizing watermark embedding to maintain effectiveness across different conditions [60, 86].

9. **Watermarking Embedding Regions.** In digital watermarking, embedding regions refer to specific areas within a digital medium, such as an image, audio, or video file, where a watermark is inserted. The selection of these regions is crucial, as it directly impacts the watermark's imperceptibility and robustness [87].

1. Entropy: Embedding region refers to selecting areas in digital content with high information content or randomness for watermark insertion, aiming to enhance robustness while minimizing impact on content quality.
2. Variance: Embedding region refers to selecting areas in digital content with high variability or contrast for watermark insertion, aiming to enhance robustness while maintaining visual quality.

10. **Watermarking Encryption Methods.** In digital watermarking, encryption methods are employed to enhance the security and robustness of the embedded watermark. By encrypting the watermark before embedding it into the host content, unauthorized detection, extraction, or tampering becomes significantly more challenging.

1. DES: In watermarking, Data Encryption Standard (DES) is a type of symmetric-key encryption, which encrypts watermark data with a secret key so that only authorized users can see the watermark [88]. Figure 12 shows the process flow of the DES algorithm.

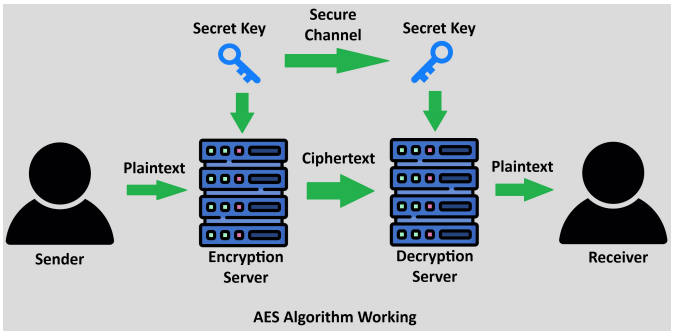

FIGURE 13. AES algorithm.

2. AES: In watermarking, Advanced Encryption Standard (AES) is a type of symmetric-key encryption, which encodes watermark data such that only authorized users with the correct key can access or extract the watermark [88]. Figure 13 shows the process flow of the AES algorithm.
3. RSA: RSA (Rivest-Shamir-Adleman) is another type of encryption that utilizes public key cryptography in order to securely encrypt the watermark information. This ensures that only the authorized users with access to the private key can decrypt it [88]. Figure 14 shows the process flow of the RSA algorithm.

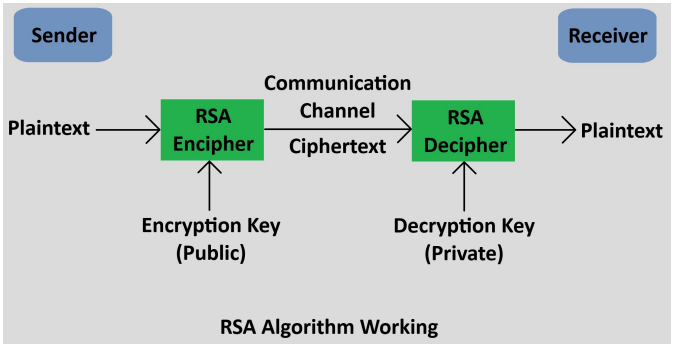

FIGURE 14. RSA algorithm.

4. Chaotic Encryption: Chaotic encryption is a cryptographic technique that utilizes chaotic dynamical systems in order to securely transmit information. This technique starts with the sender generating a binary representation of the plaintext message, which is then encrypted by constructing data blocks based on the symbolic dynamics of the chaotic systems. The resulting ciphertext consists of a series of parameters, including thresholds and initial conditions essential for the decryption process. This approach enhances security by making it difficult for unauthorized parties to decipher the message without the knowledge of the specific chaotic system and its parameters, thus providing a robust means of protecting confidential information [89].

11. **Watermarking Loss Functions.** In digital watermarking, a loss function is a mathematical measure used to evaluate the difference between the original and watermarked media, ensuring

optimal embedding and extraction. It guides the watermarking algorithm in minimizing distortions while maximizing robustness and imperceptibility. Typically, loss functions are designed to balance fidelity (how closely the watermarked content resembles the original) and robustness (resistance to attacks). Advanced watermarking techniques may also employ adversarial loss functions in deep learning-based methods to improve resilience against removal attacks. A well-optimized loss function helps achieve secure and high-quality watermarking by preserving content integrity while maintaining watermark detectability.

11.1. **Pixel-Wise Loss.** Pixel-wise loss in watermarking measures the difference between corresponding pixels of the original and watermarked images, focusing on preserving pixel-level fidelity [90].

11.2. **Perceptual Loss.** Perceptual loss in watermarking evaluates the difference between the original and watermarked images based on human perception, ensuring the watermarking process maintains visual quality and consistency [91].

11.3. **Adversarial Loss.** Adversarial loss in watermarking is used in GAN-based methods to differentiate between real and watermarked images, encouraging the watermark to be more realistic and less detectable [92].

11.4. **Prior Loss.** Prior loss in watermarking incorporates additional constraints or knowledge into the watermarking process to guide the embedding towards more desirable or expected characteristics [93].

12. **Watermarking Document Types.** In digital watermarking, document type denotes the medium embedding the watermark. Each type poses distinct challenges and demands specialized embedding techniques to ensure imperceptibility and robustness. The choice of watermarking technique depends on the characteristics of the document, with a trade-off between security, imperceptibility, and robustness against various attacks.

12.1. **Image.** Image watermarking places a watermark on digital images to protect copyright or authenticate the image without compromising image quality [94]. Figure 15a shows an example of an image watermark.

12.2. **Video.** Video watermarking applies a watermark to video content to protect intellectual property or authenticate it, and the watermark remains active during the video [95]. Figure 15b shows an example of a video watermark.

12.3. **Audio.** Audio watermarking applies a watermark to an audio signal to protect intellectual property or authenticate the signal while maintaining audio quality during embedding [96]. Figure 16 shows an example of an audio watermark.

12.4. **Text.** Text watermarking puts an authorship or authentication watermark on text, often with subtle changes to maintain readability [97]. Figure 17 shows an example of a text watermark.

13. **Deep-Learning Based Watermarking.** Deep learning-based watermarking embeds and extracts watermarks in digital media such as images and videos using neural networks. Such an approach makes the watermarking methods more robust, imperceptible, and adaptable compared to traditional methods. Deep learning-based watermarking is a promising advancement in digital media, with ongoing research addressing challenges to realize its full potential [98]. Deep Learning approach in watermarking can be classified as shown in Figure 18.

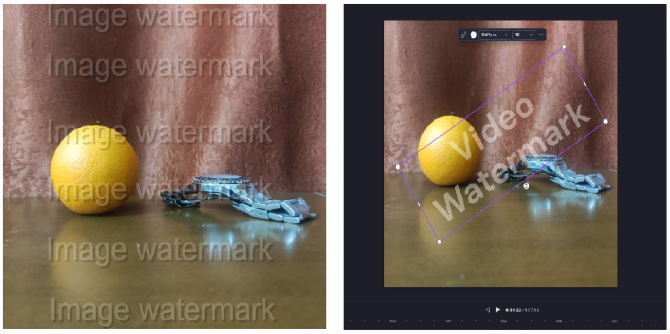

FIGURE 15. (a) Image watermark, and (b) Video watermark.

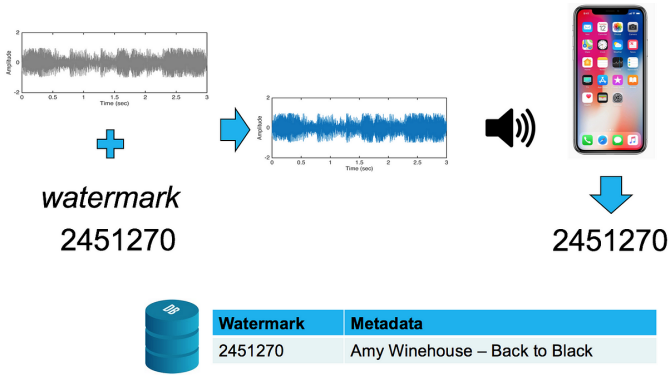

FIGURE 16. Audio watermark.

FIGURE 17. Text watermark.

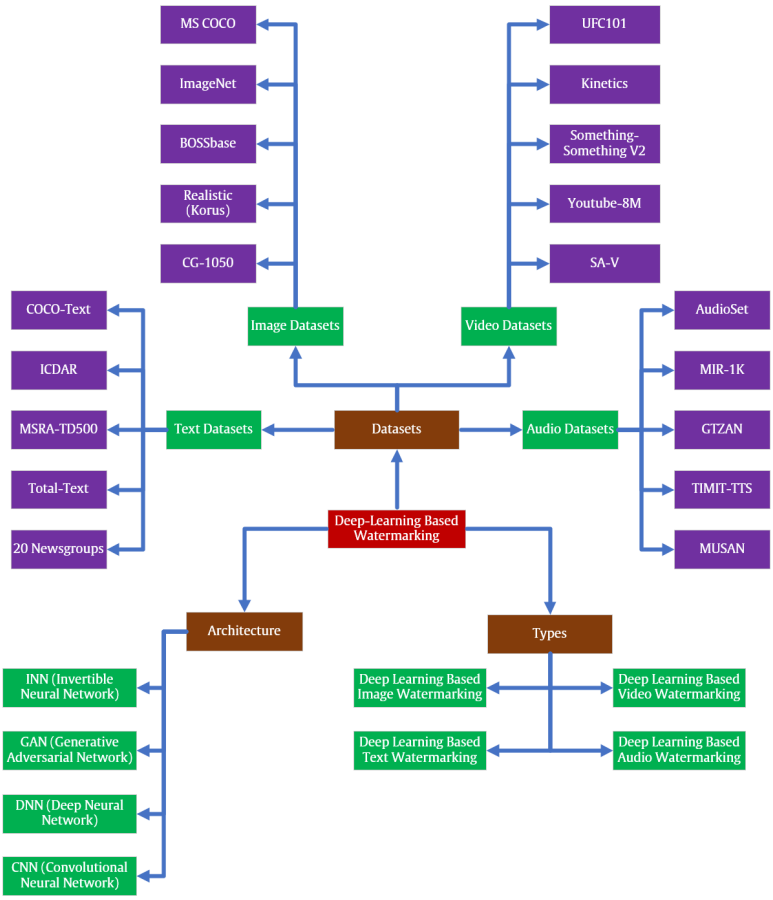

FIGURE 18. Deep learning-based watermark.

13.1. **Architecture.** Architecture in deep learning-based watermarking refers to the specific design and structure of neural networks used for embedding and detecting watermarks, such as CNNs, GANs, or DNNs. Deep learning-based watermarking techniques can be categorized into the following architectures:

13.1.1. *Deep Neural Networks (DNN).* Deep Neural Networks (DNNs) have multiple hidden layers that model complex data patterns. They are crucial in modern video and image watermarking because they learn robust, imperceptible embedding patterns, allowing watermarks to remain hidden from humans but detectable by machines even after attacks like compression, cropping, or noise. Each DNN layer extracts features from low-level edges to high-level video semantics, improving watermark robustness. However, DNN-based watermarking faces interpretability challenges, as it's hard to identify how the watermark is embedded or extracted within the network, complicating analysis of failures like watermark undetectability after attacks [99]. Figure 19 shows the DNN architecture.

13.1.2. *Convolutional Neural Networks (CNN).* Convolutional Neural Networks (CNNs) are artificial neural networks used mainly for image processing and pattern recognition. They are

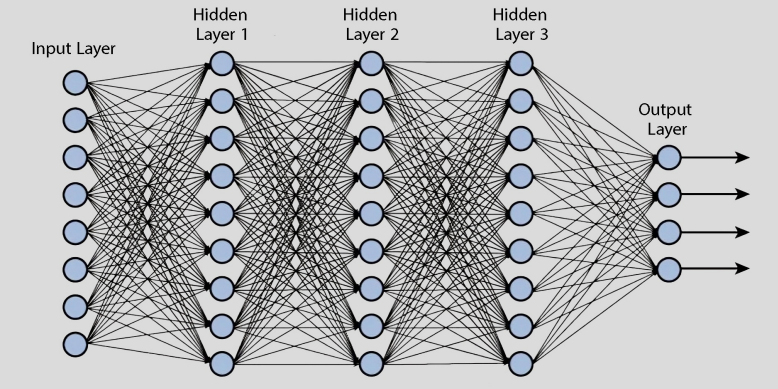

FIGURE 19. Deep neural network architecture.

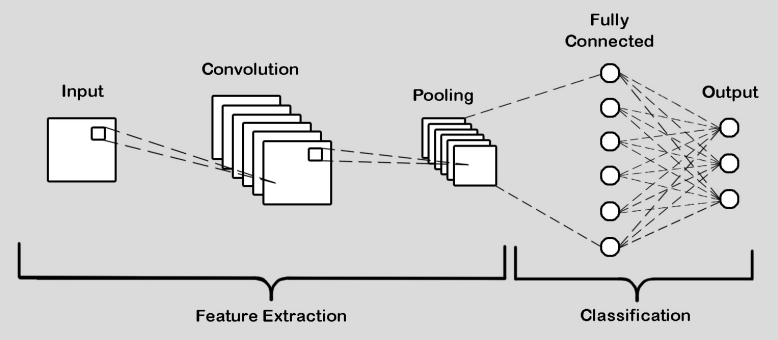

FIGURE 20. Convolutional neural network architecture.

widely applied in image and video watermarking due to their ability to learn and extract spatially coherent features, enabling imperceptible and robust watermark embedding against attacks like compression, rotation, noise, and cropping [100]. Convolutional layers use learnable filters to detect local patterns such as edges and textures, identifying low perceptual sensitivity regions for secure watermark embedding. These filters scan the image to create feature maps, guiding embedding efficiently. Weight sharing allows filters to recognize patterns anywhere, enhancing robustness. Pooling layers, especially max-pooling, improve resilience to small shifts or distortions, ensuring watermark detectability after slight content changes. Fully connected layers finalize embedding or detection decisions by interpreting extracted high-level features. Figure 20 shows the CNN architecture.

13.1.3. *Generative Adversarial Networks (GAN).* Generative Adversarial Networks (GANs) are a type of artificial intelligence algorithm designed to address the generative modeling problem. This type of network is increasingly being used in digital watermarking to enhance both the robustness and imperceptibility of watermark embedding. GANs address the generative modeling challenge by learning to create data that mimics a given distribution which in this case, they create watermarked content that is visually indistinguishable from the original, yet contains securely embedded information.

In applications such as in watermarking, GANs typically consist of two competing neural networks. They are the generator, which embeds the watermark into the original image or video

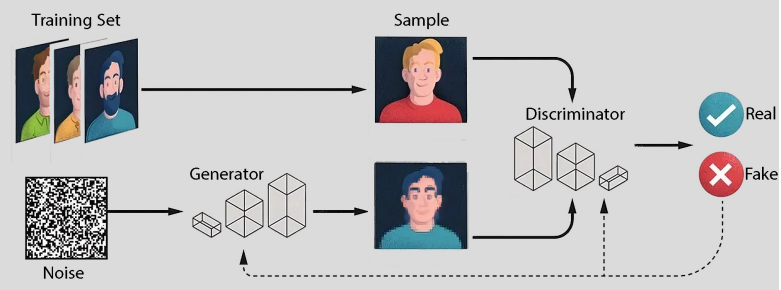

FIGURE 21. Generative adversarial network architecture.

content in a way that appears natural, and the discriminator, which attempts to distinguish between original unwatermarked and watermarked content. This adversarial training setup forms a zero-sum game, where the generator aims to fool the discriminator by producing increasingly realistic watermarked outputs. Over time, this process trains the generator to produce watermarked media that is both perceptually similar to the original and resistant to attacks such as compression, resizing, or filtering [101]. Figure 21 shows the architecture of a Generative Adversarial Network.

13.1.4. *Invertible Neural Networks (INN).* INNs are gaining attention in digital watermarking due to their unique bijective property, which allows for exact reconstruction of input data. In watermarking for example, this means that the embedding and extraction processes can be designed as perfectly reversible transformations which is an ideal scenario for maintaining image quality while ensuring reliable watermark retrieval. INNs achieve invertibility through specialized architectural components, such as coupling layers and invertible residual blocks. This ensures that each transformation that is applied to the data can be analytically or numerically reversed. This results in highly structured and controllable watermark embedding that offers precise control over where and how watermark information is embedded within the media [102]. Figure 22 shows the architecture of an Invertible Neural Network.

13.2. **Datasets.** Datasets contain digital content like images, videos, audio, or even pure text that are used to train and test watermarking algorithms for consistency across content. Datasets are divided into:

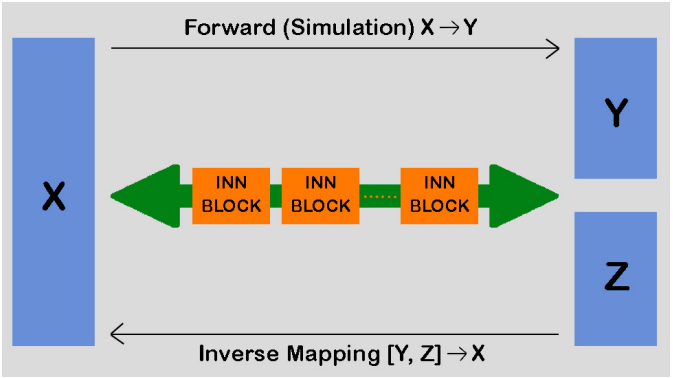

FIGURE 22. Invertible neural network architecture.

13.2.1. *Image Datasets.* One type of datasets is the image dataset which contains image files that are designed for specific purposes like training, testing, and evaluating computer vision models. These datasets usually contain metadata such as labels, annotations, or ground truth masks for classification, object detection, segmentation, and forensic analysis tasks. Image datasets may originate from real-world photographs, synthetically generated images, or a mixture thereof, and are fundamental components of artificial intelligence, image processing, and digital forensics applications [103–107].

13.2.2. *Video Datasets.* A video dataset is a structured set of video clips that are used for training, testing, and evaluating machine learning models for tasks such as action recognition, object tracking, video segmentation, and forensic analysis. These datasets can often have timestamps, object bounding boxes, motion trajectories, and ground truth labels for supervised learning. Video datasets might have real-world recordings, synthetic simulations, or a combination of both, which plays an extremely important role in computer vision, video analytics, and deep learning applications [108–112].

13.2.3. *Audio Datasets.* Audio datasets are collections of audio samples used for embedding, testing, and evaluating audio watermarking techniques. They typically contain speech, music, and environmental sounds such as audio signals to assess watermarking robustness, imperceptibility, and security [113–117].

13.2.4. *Text Datasets.* Textual datasets for watermarking are developed, tested, and evaluated in natural language processing techniques. These datasets usually contain sentences, paragraphs, or full documents as input for embedding and detecting hidden patterns or signals without losing readability or meaning. They study the robustness, imperceptibility, and security of text watermarking methods against attacks like paraphrasing, synonym substitution, and reordering [118–122].

Table 3 below, is a summary for datasets examples in each category that was explored in this sub-section.

13.3. **Types.** Deep learning-based watermarking is classified into four categories: image, video, audio, and text watermarking.

13.3.1. *Deep Learning-Based Image Watermarking.* is a modern approach to embedding and extracting hidden information (watermarks) within digital images using deep learning models. Instead of traditional algorithms, it uses the learning capabilities of neural networks to achieve imperceptibility and robustness against various attacks [123].

13.3.2. *Deep Learning-Based Video Watermarking.* Deep learning-based video watermarking is another type that is still in early stages. Its techniques can be classified based on whether they operate on the original video frames or in the compressed domain. Some methods incorporate CNNs, GANs, or attention mechanisms to improve robustness against common video-specific attacks such as frame dropping, compression, and collusion [124].

13.3.3. *Deep Learning-Based Audio Watermarking.* This is a type deep learning watermarking that involves embedding imperceptible, identifiable, and traceable watermarks into audio content using deep neural networks. This technique enhances traditional audio watermarking by improving robustness against various attacks and maintaining high audio quality. Recent advancements have focused on leveraging deep learning to achieve better performance in terms of robustness, capacity, and imperceptibility [125].

TABLE 3. Summary of Datasets.

Dataset	Category	Size	Key Metrics	Description
ImageNet	Image	Over 14M	21K categories	• Large visual database organized hierarchically for object recognition research.
MS COCO	Image	328K	2.5M labeled instances in 91 categories	• Object detection, segmentation, and captioning dataset focusing on natural, cluttered scenes.
BOSSbase	Image	Thousands	512×512 grayscale images	• Standard steganography benchmark with raw images.
Realistic (Korus)	Image	-	Pristine + tampered images from 4 DSLR models	• Evaluation of tampering localization using PRNU analysis; includes realistic and synthetic forgeries.
CG-1050	Image	1050	Tampered images from 100 originals	• Highlights altered regions; useful for training and validating tampering detection methods.
UCF101	Video	13,320 clips	101 action classes	• Real-world action recognition from YouTube videos.
Kinetics	Video	-	400 action classes with 400+ clips	• Large-scale pre-training for spatio-temporal action recognition.
YouTube-8M	Video	7M videos	4716 classes	• Video classification with large, diverse YouTube content.
Something-Something V2	Video	220,847 clips	174 labels	• Fine-grained human-object interaction dataset.
SA-V	Video	51K videos	643K segmentation masks	• Spatial-temporal segmentation for open-world video analysis.
TIMIT-TTS	Audio	80K	Artificial speech tracks	• Benchmark dataset for deepfake/synthetic speech detection.
MUSAN	Audio	109 hours	60h speech, 42h music, 6h noise	• Used for voice activity detection (VAD) and music/speech discrimination tasks.
MIR-1K	Audio	1,000 clips	4–13 seconds per clip	• Used to test source separation, lyric recognition, and music information retrieval.
GTZAN	Audio	1,000 excerpts	30s each, 10 genres	• Music genre recognition benchmark (known labeling issues).
AudioSet	Audio	1.7M clips	10s each, 632 event classes	• Hierarchical audio event recognition dataset.
COCO-Text	Text	173K annotations	63K images	• Scene text detection and recognition in natural images.
ICDAR	Text	462 images	229 train / 233 test	• Standard benchmark for near-horizontal text detection.
MSRA-TD500	Text	500 images	300 train / 200 test	• Multi-oriented text detection (Chinese & English).
Total-Text	Text	1,550 images	9,330 annotated text instances	• Advances robust text detection algorithms, especially deep learning and segmentation-based.
20 Newsgroups	Text	18K documents	20 topics	• Text classification and mining with various "bydate" splits.

13.3.4. *Deep Learning-Based Text Watermarking.* In this type, deep learning models are employed to embed and extract hidden information within text data. These models learn subtle patterns in the text that can be modified to encode a watermark without significantly altering the text's readability or meaning. The goal is to create watermarks that are imperceptible to humans but can be reliably detected by a trained deep learning model, even after various manipulations or attacks. This approach aims to improve the robustness and capacity of text watermarking compared to traditional methods [126].

14. **Watermarking Evaluation Metrics.** The performance metrics of watermarking are used to evaluate the effectiveness of watermark embedding and extraction processes, balancing key attributes like imperceptibility, robustness, and capacity. These metrics aid in optimizing watermarking techniques for a variety of applications, and ensures that the watermark meets the requirements.

These metrics often conflict with one another, which makes it challenging to optimize all aspects simultaneously. For example, increasing capacity, which is the amount of information embedded, typically reduces imperceptibility, as adding more data makes the watermark more noticeable. Similarly, enhancing robustness, which is the watermark's ability to withstand attacks, may require modifying important parts of the host media, which can also degrade visual quality. Therefore, balancing these conflicting metrics is crucial in designing effective watermarking systems, and these trade-offs must be carefully considered based on the target application. Figure 23 shows the classification of performance measures.

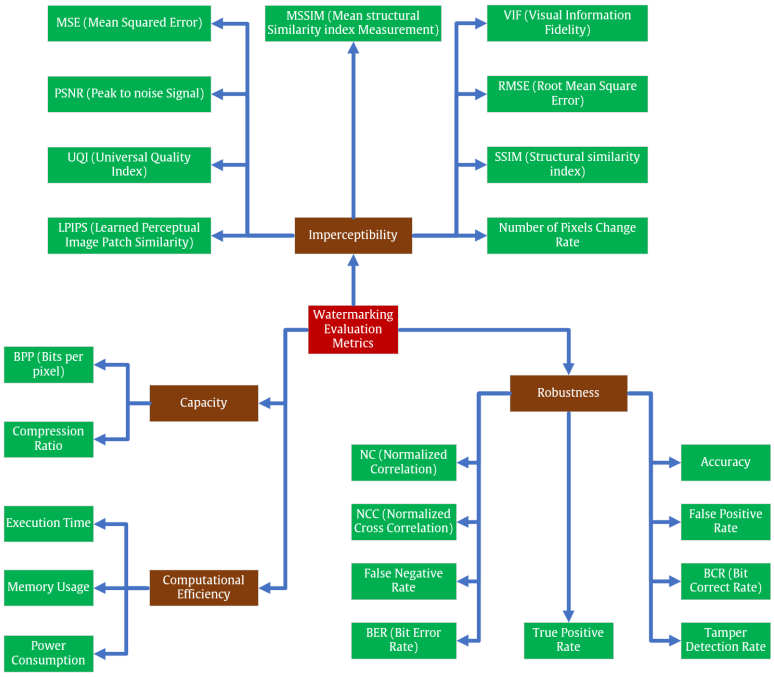

FIGURE 23. Classification of performance metrics.

14.1. **Robustness.** In this category, performance metrics for watermarking are basically measures of how well a watermark withstands attacks or distortions that are applied to the watermarked content. Such attacks may include common image processing operations like compression, resizing, noise addition, or geometric transformations like rotation and cropping. Robustness is

essential to ensure the watermark can be detected after manipulations, which is important for copyright protection and content authentication. In this category, performance metrics are:

14.1.1. *Normalized Correlation (NC).* NC is a robustness metric that compares the extracted watermark to the original embedded watermark. It measures extraction accuracy, with values closer to 1 indicating higher similarity and better quality. NC is used to evaluate robustness against attacks like compression, noise, or geometric transformations. A high NC value shows the watermark remains intact after attacks, demonstrating the watermarking method's effectiveness [29].

14.1.2. *Bit Error Rate (BER).* Bit Error Rate (BER) is another robustness performance metric in digital watermarking that measures the accuracy of the extracted watermark in comparison to the original embedded watermark. This metric measures the percentage of bits that were extracted or modified incorrectly because of watermark embedding and possible attacks. A lower BER indicates better performance, meaning fewer errors and greater watermark resistance to distortions like noise, compression, or geometric transformations [127].

14.1.3. *Bit Correct Rate (BCR).* BCR is a robustness performance metric that is utilized for digital watermarking to measure the accuracy of watermark extraction. This metric counts the number of bits that were correctly extracted out of the total number of bits in the watermark as a measure of the watermark robustness and quality after embedding and possible attacks [127].

14.1.4. *Normalized Cross Correlation (NCC).* NCC is a robustness performance metric in digital watermarking that compares the original watermark and the watermark that was extracted. This metric determines how closely the extracted watermark matches that of the original embedded watermark after possible attacks or distortions have been applied to the watermarked content. This metric is especially useful for measuring the robustness of a watermarking system under noise, compression, or geometric attacks [128].

14.1.5. *True Positive Rate.* True positive rate (TPR), also called sensitivity or recall, is a performance metric that measures the proportion of correctly detected watermarks (true positives) out of all the possible watermark instances (true positives and false negatives). A higher TPR means that the watermarking system can identify the watermark under various attacks or distortions [129].

14.1.6. *False Positive Rate (FPR).* FPR is a performance metric that measures how likely a watermark can be detected incorrectly in a content that does not contain one. Basically, it measures the probability that a watermark detection algorithm recognizes a watermark where none is embedded. A high FPR means the system frequently misclassifies unwatermarked content as watermarked, compromising watermarking reliability [130].

14.1.7. *False Negative Rate (FNR).* FNR is a performance measure that is used to measure the chance of not detecting an existing watermark in content that is embedded with it. Basically, it measures the proportion of times the watermark detector shows a watermark without one. Having a high false-positive rate would indicate that the system frequently misses the watermark detection [131].

14.1.8. *Tamper Detection Rate (TDR).* TDR is a performance metric used in digital watermarking to identify unauthorized changes to the protected content. This metric quantifies the proportion of tampered regions that were correctly detected by the system to give an idea of its sensitivity and reliability in detecting content integrity violations [132].

14.1.9. *Accuracy.* is a performance measure in digital watermarking where the detection efficiency of watermarks is measured. This metric evaluates the proportion of instances that were correctly identified (watermarked or not) of the total number of examined instances. A high accuracy indicates that the system distinguishes between content with a watermark and unmarked content [133].

14.2. **Imperceptibility.** Imperceptibility is a performance metric that is used in digital watermarking to measure how unnoticeable the embedded watermark is in the host content. An invisible watermark ensures that the quality and appearance of the original content are preserved and also virtually unchanged. This is necessary in applications where the original user experience is important, such as digital images, audio, and video. In this category, performance metrics include:

14.2.1. *Mean Squared Error (MSE).* MSE is a performance metric that is used to measure the imperceptibility of the embedded watermark in digital watermarking. This metric measures the average squared differences between corresponding pixels in the original and watermarked images, giving a quantitative evaluation of distortion that is caused by watermarking [134].

14.2.2. *Peak to noise Signal (PSNR).* PSNR is an imperceptibility performance measure of the watermark embedded in a host image. The ratio between the maximum possible power of a signal and the power of corrupting noise is expressed as a function of decibels (dB). Higher PSNR values suggest better similarity of the watermarked image to the original, which implies better imperceptibility [135].

14.2.3. *Structural similarity index (SSIM).* In digital watermarking, the SSIM is a performance metric used to assess the perceptual similarity between an original image and its watermarked version. Unlike traditional metrics such as Peak Signal-to-Noise Ratio (PSNR), which focus on pixel-wise differences, SSIM evaluates changes in structural information, luminance, and contrast to provide a measure that aligns more closely with human visual perception [136].

14.2.4. *Mean structural Similarity index Measurement (MSSIM).* In digital watermarking, the MSSIM is an imperceptibility performance metric that assesses the perceptual similarity of the original and watermarked images. This metric extends SSIM by evaluating image quality across multiple scales, offering a more comprehensive measure of image similarity [137].

14.2.5. *Root Mean Square Error (RMSE).* RMSE is another performance metric that is used to measure the imperceptibility of the embedded watermark by measuring the average error between the original image and the watermarked image. It's calculated by taking the square root of the average of the squared differences between corresponding pixels in the two images [138].

14.2.6. *Visual Information Fidelity (VIF).* VIF is a metric in watermarking that measures the perceptual quality of watermarked images by simply evaluating how well they're capable of preserving the visual information of the original content. While other traditional metrics such as Mean Squared Error (MSE) or Peak Signal to Noise Ratio (PSNR) may not be sensitive to human visual system features, VIF considers the human visual system's sensitivity to image features [139].

14.2.7. *Universal Quality Index (UQI).* In digital watermarking, UQI is a performance metric used to determine the perceptual quality of images by evaluating three components: luminance distortion, contrast distortion, and structural distortion. Unlike the more traditional metrics like Mean Squared Error (MSE) or Peak Signal-to-Noise Ratio (PSNR), which are incompatible with human visual perception, UQI provides an overall evaluation taking into account these three factors [140].

14.2.8. *Number of Pixels Change Rate (NPCR).* In digital image encryption, NPCR is a performance metric that measures the percentage of pixels in an encrypted image that differ from those in another encrypted image, which was generated by encrypting the same original image with a slight modification—typically a single pixel change. This metric assesses the sensitivity of the encryption algorithm to small changes in the plaintext image, which is crucial for evaluating the algorithm's robustness against differential attacks [141].

14.2.9. *Learned Perceptual Image Patch Similarity (LPIPS).* In digital watermarking, LPIPS is a metric that assessing perceptual similarity between the original and watermarked images. Unlike traditional metrics like PSNR or SSIM, which measure pixel-wise differences or structural information, LPIPS uses deep neural network features to assess how similar two images are to a human observer. It thus provides a more reliable measure of perceptual differences, which captures high-level semantic discrepancies not captured by standard metrics [142]. LPIPS measures the perceptual difference between the images.

14.3. **Capacity.** In digital watermarking, the capacity performance metric means how much information can be embedded in the host content, such as an image, audio, or video without compromising quality. This metric is important in applications where large amounts of data have to be embedded, such as copyright information or authentication codes, without compromising the perceptual quality of the original content. Performance metrics in this category include:

14.3.1. *Bits Per Pixel (BPP).* In digital watermarking, BPP is a performance metric which quantifies the amount of data that are embedded into an image relative to its size. This number is calculated by dividing the number of watermark bits by the total number of pixels in the image. This metric is crucial for assessing watermarking capacity, indicating how much information can be embedded without compromising image quality [143].

14.3.2. *Compression Ratio (CR).* In digital watermarking, CR is a performance measure of how efficiently a watermarking algorithm reduces data size. It calculates the ratio between the original (uncompressed) and watermarked (compressed) sizes of data. The increased compression ratio implies that the watermarking has reduced the data size for storage and transmission purposes [144].

14.4. **Computational Efficiency.** Computational efficiency in digital watermarking basically refers to how efficiently the watermarking algorithm uses computational resources such as time, memory, and power during the embedding and extraction processes. It's critical for real-time, mobile, or hardware-constrained applications. Performance metrics in this category include:

14.4.1. *Execution Time.* This metric refers to the total duration required to embed and extract the watermark using the watermark algorithm. It's used to assess the computational efficiency of the method. It reflects how quickly the algorithm can perform its operations and is usually measured in milliseconds or seconds [145].

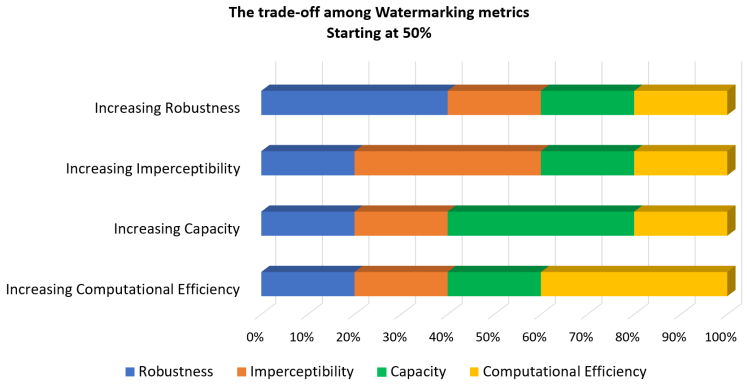

FIGURE 24. Trade-off among watermarking metrics.

14.4.2. *Memory Usage.* Memory usage is a metric that refers to the amount of memory (typically RAM) required by a watermarking algorithm during its execution, including both the embedding and extraction phases. It encompasses the memory needed to store image data, watermark data, intermediate computations, and any transformation matrices or buffers used in the process [146].

14.4.3. *Power Consumption.* Power consumption is a metric that refers to the amount of electrical energy used by a watermarking system or algorithm during its execution, particularly during the embedding and extraction processes. This metric is typically measured in watts (W) or joules (J) over time and is an important metric for evaluating the energy efficiency of the algorithm. It's critical in mobile, battery-powered, and embedded devices, where energy resources are limited [147]. Achieving a balance between these metrics (robustness, imperceptibility, capacity, and computational efficiency) in digital watermarking is important because each metric serves a different core function. And these functions often conflict with one another. Figure 24 shows a general trade-off among these watermarking metrics.

15. **Challenges.** Despite the significant advancements in digital watermarking, there are still many challenges that hinder its progress and efficiency. This section will address some of these challenges:

1. Balance Between Robustness and Imperceptibility: Balancing watermark resilience to attacks with maintaining content quality remains a challenge. For example, Adobe's open-source TrustMark watermarking system offers tunable strength levels. Its default "Quality" model is described as providing a "good trade-off between robustness and imperceptibility". Adjusting the embedding strength highlights the trade-off: increasing it (e.g., to 1.5) improves robustness against distortions but causes visible "ripple" artifacts, while lowering it enhances visual quality but reduces robustness.
2. Resistance to Complex Attacks: Watermarking systems are increasingly vulnerable to advanced attack types like desynchronization, and deepfake-based manipulations. Thus, developing robust schemes resistant to both traditional and novel attacks is crucial and challenging. For example, Zhao et al. showed in their NeurIPS paper [148] that hidden pixel-level watermarks in images are highly vulnerable to AI-based attacks. They first corrupted an image with noise, then used a diffusion model to "regenerate" it, removing 98% of the watermark in the state-of-the-art RivaGAN scheme while maintaining high image quality. In other words, a sophisticated generative (deepfake-style) process can almost completely erase the watermark [148].

3. Computational Complexity: Many watermarking algorithms, especially deep learning or hybrid ones, demand significant computational resources, limiting their practicality for real-time or resource-constrained environments. For example, Ye et al. [149] found that replacing 2D convolutions with 3D convolutions in video watermarking, to handle temporal data, caused a significant increase in computational cost. Training or embedding with large neural networks is similarly costly and such schemes need non-trivial resources [149].
4. Dynamic Content Adaptation: Multimedia content is often subject to editing, compression, or format changes. Adapting watermarking schemes to be content-aware and dynamically reconfigurable poses another layer of complexity. Zhang et al. [150] note that even subtle AI-powered edits can destroy embedded marks. They report that "watermark corruption after editing is primarily caused by VAE compression" when an image is auto-reencoded, so they explicitly train for global editing robustness. In practice, this means a watermark invisible in the original may disappear after standard processing like social-media recompression or format conversion [150].
5. Interpretability in Deep Learning-Based Approaches: While deep learning has shown promise, the "black box" nature of neural networks complicates the understanding and debugging of watermarking behavior, especially during extraction and error analysis. For example, Xue et al. [151] explicitly state that prior DNN watermarking techniques "need to modify the model and lack interpretability". The need for an "interpretable" method shows that without it, it's unclear what features carry the watermark. Most DL watermark schemes act as black boxes, highlighting the interpretability challenge [151].
6. Standardization Issues: The lack of universal standards for watermarking techniques, metrics, and testing leads to inconsistent evaluation and deployment. As early as 1999, Kutter & Petitcolas pointed out that a number of papers on "robust" digital watermarking systems have been presented but "none of them uses the same robustness criteria". They wrote that papers used different measurement methods, making comparison "not practical at all and slows down progress" [152]. This fragmentation persists, prompting efforts like the WAVES benchmark, which aims to "overcome the limitations of current evaluation methods" through a unified protocol and robustness metrics. The need for WAVES itself illustrates the prior lack of agreed-upon benchmarks or metrics in watermarking.

16. **Conclusion.** This literature review has explored the concepts, classifications, methodologies, and advancements in the field of digital watermarking. From traditional spatial and transform domains to advanced deep learning-based techniques. The review highlighted the evolving landscape of watermarking solutions that are tailored to meet the growing digital security needs. Additionally, optimization techniques, attack resistance strategies, and performance metrics were examined to provide a thorough understanding of the current systems. A particular focus was placed on the integration of deep learning in this field, especially of CNNs and GANs, which has pushed the boundaries of watermarking in terms of imperceptibility and robustness. Furthermore, the paper outlined the core challenges such as computational cost, attack resilience, and interpretability that continue to shape future research directions. By addressing these challenges and improving scalability, standardization, and computational costs, digital watermarking can serve as a cornerstone technology for safeguarding digital media in an increasingly interconnected and data-driven world. Ultimately, this paper served as an essential resource for understanding watermarking aspects and serves as a guiding resource for researchers, developers, and practitioners that aim to design more secure, efficient, and intelligent watermarking frameworks.

Acknowledgment. The authors would like to thank University of Baghdad for their general support. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] N. Safizadeh and S. H. R. Ahmadi, "A comprehensive analysis of the false-positive problem in svd-based image watermarking," *Multimedia Tools and Applications*, vol. 82, no. 14, pp. 22 275–22 295, 2023.

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

[3] B. Madhushree, H. Basanth Kumar, and H. Chennamma, "An exhaustive review of authentication, tamper detection with localization and recovery techniques for medical images," *Multimedia Tools and Applications*, vol. 83, no. 13, pp. 39 779–39 821, 2024.

[4] M. Ansarian and Z. Baharlouei, "Applications and challenges of telemedicine: privacy-preservation as a case study," *Archives of Iranian medicine*, vol. 26, no. 11, p. 654, 2023.

[5] F. Ernawan and D. Ariatmanto, "A recent survey on image watermarking using scaling factor techniques for copyright protection," *Multimedia Tools and Applications*, vol. 82, no. 18, pp. 27 123–27 163, 2023.

[6] H. J. Kim and N.-K. Baik, "A research on information security based on digital watermarking," *Asia-pacific Journal of Convergent Research Interchange (APJCRI)*, pp. 1–15, 2023.

[7] Y. Cui, J. Ren, H. Xu, P. He, H. Liu, L. Sun, Y. Xing, and J. Tang, "Diffusionshield: A watermark for copyright protection against generative diffusion models," *arXiv preprint arXiv:2306.04642*, 2023.

[8] M. S. Moad, M. R. Kafi, and A. Khaldi, "A wavelet based medical image watermarking scheme for secure transmission in telemedicine applications," *Microprocessors and Microsystems*, vol. 90, p. 104490, 2022.

[9] Y. Zhang, Q. Ding, B. Gao, F. Zhang, and X. Yang, "Robust digital watermarking for remote sensing images using generative adversarial networks and adaptive channel attention mechanisms," in *Proceedings of the 2024 2nd International Conference on Artificial Intelligence, Systems and Network Security*, 2024, pp. 89–94.

[10] S. Emmanuel, A. Vinod, D. Rajan, and C. Heng, "An authentication watermarking scheme with transaction tracking enabled," in *2007 Inaugural IEEE-IES Digital EcoSystems and Technologies Conference*. IEEE, 2007, pp. 481–486.

[11] M. AlShaikh, "A novel tamper detection watermarking approach for improving image integrity," *Multimedia Tools and Applications*, vol. 82, no. 7, pp. 10 039–10 060, 2023.

[12] P. Garg and A. Jain, "A robust technique for biometric image authentication using invisible watermarking," *Multimedia Tools and Applications*, vol. 82, no. 2, pp. 2237–2253, 2023.

[13] D. Raj and D. Sharma, "Enhancing digital image forensics and security: Innovations in metadata, watermarking and blockchain technology," in *2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0*. IEEE, 2024, pp. 1–6.

[14] M. Li and Y. Yue, "Security analysis and improvement of dual watermarking framework for multimedia privacy protection and content authentication," *Mathematics*, vol. 11, no. 7, p. 1689, 2023.

[15] K.-K. Tseng, Q. Na, and R. F.-Y. Lin, "A document management system with multi-biometric watermarking for educational purpose," in *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. Springer, 2022, pp. 265–275.

[16] A. Agrawal, K. Sethi, and P. Bera, "Blockchain-based cardinal e-voting system using biometrics, watermarked qr code and partial homomorphic encryption," in *Proceedings of the International Conference on Cybersecurity, Situational Awareness and Social Media: Cyber Science 2022; 20–21 June; Wales*. Springer, 2023, pp. 411–436.

[17] H. Chaudhary, P. Garg, and V. P. Vishwakarma, "Enhanced medical image watermarking using hybrid dwt-hmd-svd and arnold scrambling," *Scientific Reports*, vol. 15, no. 1, p. 9710, 2025.

[18] J. Fairoze, S. Garg, S. Jha, S. Mahloujifar, M. Mahmoody, and M. Wang, "Publicly-detectable watermarking for language models," *arXiv preprint arXiv:2310.18491*, 2023.

[19] W. Cheung, "Digital image watermarking in spatial and transform domains," *2000 TENCON proceedings. Intelligent systems and Technologies for the new Millennium (cat. No. 00CH37119)*, vol. 3, pp. 374–378, 2000.

[20] R. K. Singh, A. P. Dube, and R. Singh, "Least significant bit-based image watermarking mechanism: a review," *International Journal of Social Ecology and Sustainable Development (IJSESD)*, vol. 13, no. 1, pp. 1–9, 2022.

[21] T. K. Das, "Cryptanalysis of block based spatial domain watermarking schemes," in *Progress in Cryptology-INDOCRYPT 2003: 4th International Conference on Cryptology in India, New Delhi, India, December 8-10, 2003. Proceedings 4*. Springer, 2003, pp. 363–374.

[22] Z. B. Faheem, A. Ishaq, F. Rustam, I. de la Torre Díez, D. Gavilanes, M. M. Vergara, and I. Ashraf, "Image watermarking using least significant bit and canny edge detection," *Sensors*, vol. 23, no. 3, p. 1210, 2023.

[23] A. Tareef, K. Al-Tarawneh, and A. Sleit, "Block-based watermarking for robust authentication and integration of gis data," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 16 340–16 345, 2024.

[24] Z. Yuan, Q. Su, D. Liu, X. Zhang, and T. Yao, "Fast and robust image watermarking method in the spatial domain," *IET Image Processing*, vol. 14, no. 15, pp. 3829–3838, 2020.

[25] L. Laouamer, "A new image watermarking technique in spatial domain using dc coefficients and graph representation," in *International Conference on Advanced Machine Learning Technologies and Applications*. Springer, 2019, pp. 633–644.

[26] L. Liu, B. Yang, P. Fieguth, Z. Yang, and Y. Wei, "Brint: A binary rotation invariant and noise tolerant texture descriptor," in *2013 IEEE International Conference on Image Processing*. IEEE, 2013, pp. 255–259.

[27] L. AlShehri, M. Hussain, H. Aboalsamh, and A. Wadood, "Fragile watermarking for image authentication using brint and elm," *Multimedia Tools and Applications*, vol. 79, pp. 29 199–29 223, 2020.

[28] T. Pardhu and B. R. Perli, "Digital image watermarking in frequency domain," in *2016 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2016, pp. 0208–0211.

[29] W. Alomoush, O. A. Khashan, A. Alrosan, H. H. Attar, A. Almomani, F. Alhosban, and S. N. Makhadmeh, "Digital image watermarking using discrete cosine transformation based linear modulation," *Journal of Cloud Computing*, vol. 12, no. 1, p. 96, 2023.

[30] M. Begum, S. B. Shorif, M. S. Uddin, J. Ferdush, T. Jan, A. Barros, and M. Whaiduzzaman, "Image watermarking using discrete wavelet transform and singular value decomposition for enhanced imperceptibility and robustness," *Algorithms*, vol. 17, no. 1, p. 32, 2024.

[31] M. Narang and S. Vashisth, "Digital watermarking using discrete wavelet transform," *International Journal of Computer Applications*, vol. 74, no. 20, 2013.

[32] N. Zermi, A. Khaldi, M. R. Kafi, F. Kahlessenane, and S. Euschi, "A lossless dwt-svd domain watermarking for medical information security," *Multimedia Tools and Applications*, vol. 80, pp. 24 823–24 841, 2021.

[33] A. Khaldi, "Steganographic techniques classification according to image format," *International Annals of Science*, vol. 8, no. 1, pp. 143–149, 2020.

[34] A. T. S. Ho, X. Zhu, and J. Shen, "Slant transform watermarking for digital images," in *Visual Communications and Image Processing 2003*, vol. 5150. SPIE, 2003, pp. 1912–1920.

[35] A. A.-A. Hadad, H. N. Khalid, Z. S. Naser, and M. S. Taha, "A robust color image watermarking scheme based on discrete wavelet transform domain and discrete slantlet transform technique," *Ingenierie des Systemes d'Information*, vol. 27, no. 2, p. 313, 2022.

[36] J. W. Cooley, P. A. W. Lewis, and P. D. Welch, "The fast fourier transform and its applications," *IEEE Transactions on Education*, vol. 12, no. 1, pp. 27–34, 1969.

[37] A. Dey, P. Chowdhuri, and P. Pal, "Integer wavelet transform based watermarking scheme for medical image authentication," *Multimedia Tools and Applications*, vol. 83, no. 32, pp. 78 001–78 022, 2024.

[38] M. T. Taba, "The fractional fourier transform and its application to digital watermarking," in *2013 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA)*. IEEE, 2013, pp. 262–266.

[39] J.-G. Cao, J. E. Fowler, and N. H. Younan, "An image-adaptive watermark based on a redundant wavelet transform," in *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*, vol. 2. IEEE, 2001, pp. 277–280.

[40] S. Gaur and V. K. Srivastava, "A rdwt and block-svd based dual watermarking scheme for digital images," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 4, pp. 211–219, 2017.

[41] K.-L. Hua, B.-R. Dai, K. Srinivasan, Y.-H. Hsu, and V. Sharma, "A hybrid nsct domain image watermarking scheme," *EURASIP Journal on Image and Video Processing*, vol. 2017, pp. 1–17, 2017.

[42] S. Thakur, A. K. Singh, and S. P. Ghrera, "Nsct domain–based secure multiple-watermarking technique through lightweight encryption for medical images," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 2, p. e5108, 2021.

[43] I. A. Ismail, M. A. Shouman, K. M. Hosny, and H. M. Salam, "Invariant image watermarking using accurate zernike moments," *Journal of computer Science*, vol. 6, no. 1, p. 52, 2010.

[44] K. M. Hosny, "Fast computation of accurate zernike moments," *Journal of Real-Time Image Processing*, vol. 3, pp. 97–107, 2008.

[45] B. M. Mahmmod, S. H. Abdulhussain, T. Suk, M. Alsabah, and A. Hussain, "Accelerated and improved stabilization for high order moments of racah polynomials," *IEEE Access*, vol. 11, pp. 110 502–110 521, 2023.

[46] M. El Mallahi, A. Zouhri, A. Mesbah, A. Berrahou, I. El Affar, and H. Qjidaa, "Radial invariant of 2d and 3d racah moments," *Multimedia Tools and Applications*, vol. 77, pp. 6583–6604, 2018.

[47] H. S. Radeaf, B. M. Mahmmod, S. H. Abdulhussain, and D. Al-Jumaeily, "A steganography based on orthogonal moments," in *Proceedings of the International Conference on Information and Communication Technology*, ser. ICICT '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 147–153.

[48] S. H. Abdulhussain, A. R. Ramli, A. J. Hussain, B. M. Mahmmod, and W. A. Jassim, "Orthogonal polynomial embedded image kernel," in *Proceedings of the International Conference on Information and Communication Technology*, ser. ICICT '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 215–221.

[49] R. Mukundan, S. H. Ong, and P. A. Lee, "Image analysis by tchebichef moments," *IEEE Transactions on image Processing*, vol. 10, no. 9, pp. 1357–1364, 2001.

[50] P.-T. Yap, R. Paramesran, and S.-H. Ong, "Image analysis by krawtchouk moments," *IEEE Transactions on image processing*, vol. 12, no. 11, pp. 1367–1377, 2003.

[51] M. Begum, J. Ferdush, and M. S. Uddin, "A hybrid robust watermarking system based on discrete cosine transform, discrete wavelet transform, and singular value decomposition," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 8, pp. 5856–5867, 2022.

[52] R. Srivastava, R. Tomar, M. Gupta, A. K. Yadav, and J. Park, "Image watermarking approach using a hybrid domain based on performance parameter analysis," *Information*, vol. 12, no. 8, p. 310, 2021.

[53] E. Elbasi, N. Mostafa, and E. Cina, "Robust, secure and semi-blind watermarking technique using flexible scaling factor in block-based wavelet algorithm," *Electronics*, vol. 11, no. 22, p. 3680, 2022.

[54] K. M. Hosny and M. M. Darwish, "Robust color image watermarking using multiple fractional-order moments and chaotic map," *Multimedia Tools and Applications*, vol. 81, no. 17, pp. 24 347–24 375, 2022.

[55] Z. Yin, H. Yin, H. Su, X. Zhang, and Z. Gao, "Decision-based iterative fragile watermarking for model integrity verification," *arXiv preprint arXiv:2305.09684*, 2023.

[56] A. Senol, E. Elbasi, A. E. Topcu, and N. Mostafa, "A semi-fragile, inner-outer block-based watermarking method using scrambling and frequency domain algorithms," *Electronics*, vol. 12, no. 4, p. 1065, 2023.

[57] R. Ma, M. Guo, Y. Hou, F. Yang, Y. Li, H. Jia, and X. Xie, "Towards blind watermarking: Combining invertible and non-invertible mechanisms," in *Proceedings of the 30th ACM International Conference on Multimedia*, 2022, pp. 1532–1542.

[58] L. Laouamer and M. Alswailim, "Local entropy-based non blind robust image watermarking: Case of medical images."

[59] C.-F. Lee, Z.-C. Chao, J.-J. Shen, and A. U. Rehman, "A robust semi-blind watermarking technology for resisting jpeg compression based on deep convolutional generative adversarial networks," *Symmetry*, vol. 17, no. 1, p. 98, 2025.

[60] O. Juarez-Sandoval, E. Fragoso-Navarro, M. Cedillo-Hernandez, M. Nakano, H. Perez-Meana, and A. Cedillo-Hernandez, "Improved unseen-visible watermarking for copyrigth protection of digital image," in *2017 5th International Workshop on Biometrics and Forensics (IWBF)*. IEEE, 2017, pp. 1–5.

[61] S. Voloshynovskiy, S. Pereira, T. Pun, J. J. Eggers, and J. K. Su, "Attacks on digital watermarks: classification, estimation based attacks, and benchmarks," *IEEE communications Magazine*, vol. 39, no. 8, pp. 118–126, 2001.

[62] K. Srinivas and D. R. Rani, "Digital image watermarking: Framework, optimization and ai for enhanced security," *African Journal of Biological Science*, vol. 6, no. 14, pp. 5093–5119, 2024.

[63] D. Liu, D. Liu, B. Wang, and P. Zheng, "Hybrid domain digital watermarking scheme based on improved differential evolution algorithm and singular value block embedding," *IET Image Processing*, vol. 17, no. 8, pp. 2516–2536, 2023.

[64] H. Rezaei, O. Bozorg-Haddad, and X. Chu, "Grey wolf optimization (gwo) algorithm," in *Advanced optimization by nature-inspired algorithms*. Springer, 2017, pp. 81–91.

[65] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*, vol. 4. ieee, 1995, pp. 1942–1948.

[66] M. Yamni, A. Daoui, H. Karmouni, M. Sayyouri, H. Qjidaa, C. Wang, and M. O. Jamil, "A powerful zero-watermarking algorithm for copyright protection of color images based on

quaternion radial fractional hahn moments and artificial bee colony algorithm," *Circuits, Systems, and Signal Processing*, vol. 42, no. 9, pp. 5602–5633, 2023.

[67] D. K. Mahto and A. K. Singh, "Firefly optimization-based dual watermarking for colour images with improved capacity," *Multimedia Tools and Applications*, vol. 83, no. 22, pp. 61 539–61 557, 2024.

[68] M. Gupta and R. R. Kishore, "Robust digital image watermarking using cuckoo search optimization and probabilistic neural network," *Multimedia Tools and Applications*, vol. 83, no. 16, pp. 46 825–46 850, 2024.

[69] X.-S. Yang and A. Hossein Gandomi, "Bat algorithm: a novel approach for global engineering optimization," *Engineering computations*, vol. 29, no. 5, pp. 464–483, 2012.

[70] W. Wang and L. Feng, "Robust image watermarking using ant colony optimization and fast generic radial harmonic fourier moment calculation," *IET Image Processing*, vol. 18, no. 5, pp. 1200–1212, 2024.

[71] C. Zhao, L. Wu, C. Zuo, and H. Zhang, "An improved fruit fly optimization algorithm with q-learning for solving distributed permutation flow shop scheduling problems," *Complex & Intelligent Systems*, vol. 10, no. 5, pp. 5965–5988, 2024.

[72] I. Al-Aiash, R. Alquran, M. AlJamal, A. Alsarhan, M. Aljaidi, and D. Al-Fraihat, "Optimized digital watermarking: Harnessing the synergies of schur matrix factorization, dct, and dwt for superior image ownership proofing," *Multimedia Tools and Applications*, pp. 1–36, 2024.

[73] M. Agoyi, E. Çelebi, and G. Anbarjafari, "A watermarking algorithm based on chirp z-transform, discrete wavelet transform, and singular value decomposition," *Signal, Image and Video Processing*, vol. 9, pp. 735–745, 2015.

[74] C.-H. Lin, C.-Y. Lee, S.-Y. Lu, and S.-P. Chien, "Unseen visible watermarking for gray level images based on gamma correction," in *International Conference on Future Generation Communication and Networking*. Springer, 2011, pp. 236–243.

[75] W. Wenying, H. Jingxin, and Y. Yi, "Identification of blurred extent with rotation motion blurred image," in *Proceedings of the 2008 International Symposium on Information Science and Engineering*, 2008, pp. 669–672.

[76] Z.-x. Wang and X.-c. Meng, "Digital image information hiding algorithm research based on ldpc code," *EURASIP Journal on Image and Video Processing*, vol. 2018, pp. 1–12, 2018.

[77] B. D. Majeed, A. H. Taherinia, H. S. Yazdi, and A. Harati, "Csrwa: Covert and severe attacks resistant watermarking algorithm." *Computers, Materials & Continua*, vol. 82, no. 1, 2025.

[78] A. E. Aydemir, A. Temizel, and T. T. Temizel, "The effects of jpeg and jpeg2000 compression on attacks using adversarial examples," *arXiv preprint arXiv:1803.10418*, 2018.

[79] A. Shafique, J. Ahmed, M. U. Rehman, and M. M. Hazzazi, "Noise-resistant image encryption scheme for medical images in the chaos and wavelet domain," *IEEE Access*, vol. 9, pp. 59 108–59 130, 2021.

[80] J. Ebrahimnejad, A. Naghsh, and H. Pourghasem, "A robust watermarking approach against high-density salt and pepper noise (rwspn) to enhance medical image security," *IET Image Processing*, vol. 18, no. 1, pp. 116–128, 2024.

[81] S. Sarkar, A. R. Babu, S. Mousavi, S. Ghorbanpour, V. Gundecha, A. Guillen, R. Luna, and A. Naug, "Robustness with query-efficient adversarial attack using reinforcement learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 2330–2337.

[82] B. Peng, B. Peng, J. Zhou, J. Xia, and L. Liu, "Speckle-variant attack: Toward transferable adversarial attack to sar target recognition," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[83] S. W. Hasinoff, "Photon, poisson noise," in *Computer vision*. Springer, 2014, pp. 608–610.

[84] H. K. Verma, A. N. Singh, and R. Kumar, "Robustness of the digital image watermarking techniques against brightness and rotation attack," *arXiv preprint arXiv:0909.3554*, 2009.

[85] F. Zhang, H. Wang, M. He, and J. Xia, "Robust blind symmetry-based watermarking in the frequency domain against social network processing and desynchronization attacks," *IEEE Transactions on Circuits and Systems for Video Technology*, 2024.

[86] F. Ernawan and D. Ariatmanto, "An efficient adaptive scaling factor for 4× 4 dct image watermarking," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 8603–8621, 2023.

[87] D. Ariatmanto and F. Ernawan, "An improved robust image watermarking by using different embedding strengths," *Multimedia Tools and Applications*, vol. 79, no. 17, pp. 12 041–12 067, 2020.

[88] P. Parikh, N. Patel, D. Patel, P. Modi, and H. Kaur, "Ciphering the modern world: A comprehensive analysis of des, aes, rsa and dhke," in *2024 11th International Conference on Computing for Sustainable Global Development (INDIACom)*. IEEE, 2024, pp. 838–842.

[89] E. Alvarez, A. Fernández, P. Garcıa, J. Jiménez, and A. Marcano, "New approach to chaotic encryption," *Physics Letters A*, vol. 263, no. 4-6, pp. 373–375, 1999.

[90] H. Kweon, H. Kim, Y. Kang, Y. Yoon, W. Jeong, and K.-J. Yoon, "Pixel-wise deep image stitching," *arXiv preprint arXiv:2112.06171*, 2021.

[91] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, "Low-dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1348–1357, 2018.

[92] C. Wang, C. Xu, C. Wang, and D. Tao, "Perceptual adversarial networks for image-to-image transformation," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 4066–4079, 2018.

[93] R. El Jurdi, C. Petitjean, P. Honeine, V. Cheplygina, and F. Abdallah, "High-level prior-based loss functions for medical image segmentation: A survey," *Computer Vision and Image Understanding*, vol. 210, p. 103248, 2021.

[94] Z. Jiang, M. Guo, Y. Hu, J. Jia, and N. Z. Gong, "Certifiably robust image watermark," in *European Conference on Computer Vision*. Springer, 2024, pp. 427–443.

[95] M. He, H. Wang, F. Zhang, S. M. Abdullahi, and L. Yang, "Robust blind video watermarking against geometric deformations and online video sharing platform processing," *IEEE Transactions on Dependable and Secure Computing*, vol. 20, no. 6, pp. 4702–4718, 2022.

[96] M. Charfeddine, E. Mezghani, S. Masmoudi, C. B. Amar, and H. Alhumyani, "Audio watermarking for security and non-security applications," *IEEE Access*, vol. 10, pp. 12 654–12 677, 2022.

[97] Y. Lu, A. Liu, D. Yu, J. Li, and I. King, "An entropy-based text watermarking detection method," *arXiv preprint arXiv:2403.13485*, 2024.

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

[99] W. Ding, Y. Ming, Z. Cao, and C.-T. Lin, "A generalized deep neural network approach for digital watermarking analysis," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 3, pp. 613–627, 2021.

[100] K. O'shea and R. Nash, "An introduction to convolutional neural networks," *arXiv preprint arXiv:1511.08458*, 2015.

[101] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[102] Z. He, R. Hu, J. Wu, T. Luo, and H. Xu, "A transformer-based invertible neural network for robust image watermarking," *Journal of Visual Communication and Image Representation*, vol. 104, p. 104317, 2024.

[103] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, pp. 211–252, 2015.

[104] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer vision–ECCV 2014: 13th European conference, zurich, Switzerland, September 6-12, 2014, proceedings, part v 13*. Springer, 2014, pp. 740–755.

[105] V. Sedighi, J. Fridrich, and R. Cogranne, "Toss that bossbase, alice!" in *IS&T intl symposium on Electronic Imaging*, 2016.

[106] P. Korus and J. Huang, "Multi-scale analysis strategies in prnu-based tampering localization," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 4, pp. 809–824, 2016.

[107] M. Castro, D. M. Ballesteros, and D. Renza, "A dataset of 1050-tampered color and grayscale images (cg-1050)," *Data in brief*, vol. 28, p. 104864, 2020.

[108] K. Soomro, A. R. Zamir, and M. Shah, "Ucf101: A dataset of 101 human actions classes from videos in the wild," *arXiv preprint arXiv:1212.0402*, 2012.

[109] J. Carreira and A. Zisserman, "Quo vadis, action recognition? a new model and the kinetics dataset," in *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 6299–6308.

[110] S. Abu-El-Haija, N. Kothari, J. Lee, P. Natsev, G. Toderici, B. Varadarajan, and S. Vijayanarasimhan, "Youtube-8m: A large-scale video classification benchmark," *arXiv preprint arXiv:1609.08675*, 2016.

[111] R. Goyal, S. Ebrahimi Kahou, V. Michalski, J. Materzynska, S. Westphal, H. Kim, V. Haenel, I. Fruend, P. Yianilos, M. Mueller-Freitag *et al.*, "The" something something" video database for learning and evaluating visual common sense," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 5842–5850.

[112] A. Lou, Y. Li, Y. Zhang, R. F. Labadie, and J. Noble, "Zero-shot surgical tool segmentation in monocular video using segment anything model 2," in *Medical Imaging 2025: Image Processing*, vol. 13406. SPIE, 2025, pp. 718–723.

[113] D. Salvi, B. Hosler, P. Bestagini, M. C. Stamm, and S. Tubaro, "Timit-tts: A text-to-speech dataset for multimodal synthetic media detection," *IEEE access*, vol. 11, pp. 50 851–50 866, 2023.

[114] D. Snyder, G. Chen, and D. Povey, "Musan: A music, speech, and noise corpus," *arXiv preprint arXiv:1510.08484*, 2015.

[115] C.-L. Hsu and J.-S. R. Jang, "On the improvement of singing voice separation for monaural recordings using the mir-1k dataset," *IEEE transactions on audio, speech, and language processing*, vol. 18, no. 2, pp. 310–319, 2009.

[116] B. L. Sturm, "The gtzan dataset: Its contents, its faults, their effects on evaluation, and its future use," *arXiv preprint arXiv:1306.1461*, 2013.

[117] J. F. Gemmeke, D. P. Ellis, D. Freedman, A. Jansen, W. Lawrence, R. C. Moore, M. Plakal, and M. Ritter, "Audio set: An ontology and human-labeled dataset for audio events," in *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2017, pp. 776–780.

[118] A. Veit, T. Matera, L. Neumann, J. Matas, and S. Belongie, "Coco-text: Dataset and benchmark for text detection and recognition in natural images," *arXiv preprint arXiv:1601.07140*, 2016.

[119] D. Karatzas, F. Shafait, S. Uchida, M. Iwamura, L. G. i Bigorda, S. R. Mestre, J. Mas, D. F. Mota, J. A. Almazan, and L. P. De Las Heras, "Icdar 2013 robust reading competition," in *2013 12th international conference on document analysis and recognition*. IEEE, 2013, pp. 1484–1493.

[120] C. Yao, X. Bai, W. Liu, Y. Ma, and Z. Tu, "Detecting texts of arbitrary orientations in natural images," in *2012 IEEE conference on computer vision and pattern recognition*. IEEE, 2012, pp. 1083–1090.

[121] C. K. Ch'ng and C. S. Chan, "Total-text: A comprehensive dataset for scene text detection and recognition," in *2017 14th IAPR international conference on document analysis and recognition (ICDAR)*, vol. 1. IEEE, 2017, pp. 935–942.

[122] K. Albishre, M. Albathan, and Y. Li, "Effective 20 newsgroups dataset cleaning," in *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, vol. 3. IEEE, 2015, pp. 98–101.

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

[124] S. Mansour, S. B. Jabra, and E. Zagrouba, "A robust deep learning-based video watermarking using mosaic generation." in *VISIGRAPP (5: VISAPP)*, 2023, pp. 668–675.

[125] M. K. Singh, N. Takahashi, W. Liao, and Y. Mitsufuji, "Silentcipher: Deep audio watermarking," *arXiv preprint arXiv*, vol. 2406, p. 03822, 2024.

[126] T. Munyer and X. Zhong, "Deeptextmark: Deep learning based text watermarking for detection of large language model generated text," *arXiv e-prints*, pp. arXiv–2305, 2023.

[127] D. A. Basnayaka and J. Jia, "Bit error rate performance and diversity analysis for mediumband wireless communication," in *2023 IEEE Virtual Conference on Communications (VCC)*. IEEE, 2023, pp. 224–229.

[128] D. Buniatyan, T. Macrina, D. Ih, J. Zung, and H. S. Seung, "Deep learning improves template matching by normalized cross correlation," *arXiv preprint arXiv:1705.08593*, 2017.

[129] H. Holzmann and B. Klar, "Robust performance metrics for imbalanced classification problems," *arXiv preprint arXiv:2404.07661*, 2024.

[130] M. Singer-Berk, S. Gudmundsson, S. Baxter, E. G. Seaby, E. England, J. C. Wood, R. G. Son, N. A. Watts, K. J. Karczewski, S. M. Harrison *et al.*, "Advanced variant classification framework reduces the false positive rate of predicted loss-of-function variants in population sequencing data," *The American Journal of Human Genetics*, vol. 110, no. 9, pp. 1496–1508, 2023.

[131] J. N. Kanji, N. Zelyas, C. MacDonald, K. Pabbaraju, M. N. Khan, A. Prasad, J. Hu, M. Diggle, B. M. Berenger, and G. Tipples, "False negative rate of covid-19 pcr testing: a discordant testing analysis," *Virology journal*, vol. 18, no. 1, pp. 1–6, 2021.

[132] S. Gull, R. F. Mansour, N. O. Aljehane, and S. A. Parah, "A self-embedding technique for tamper detection and localization of medical images for smart-health," *Multimedia Tools and Applications*, vol. 80, pp. 29 939–29 964, 2021.

[133] C. W. Fisher, E. J. Lauria, and C. C. Matheus, "An accuracy metric: Percentages, randomness, and probabilities," *Journal of Data and Information Quality (JDIQ)*, vol. 1, no. 3, pp. 1–21, 2009.

[134] J. Ren, M. Zhang, C. Yu, and Z. Liu, "Balanced mse for imbalanced visual regression," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 7926–7935.

[135] O. Keleş, M. A. Yılmaz, A. M. Tekalp, C. Korkmaz, and Z. Doğan, "On the computation of psnr for a set of images or video," in *2021 Picture Coding Symposium (PCS)*. IEEE, 2021, pp. 1–5.

[136] D. R. I. M. Setiadi, "Psnr vs ssim: imperceptibility quality assessment for image steganography," *Multimedia Tools and Applications*, vol. 80, no. 6, pp. 8423–8444, 2021.

[137] R. Dosselmann and X. D. Yang, "A comprehensive assessment of the structural similarity index," *Signal, Image and Video Processing*, vol. 5, pp. 81–91, 2011.

[138] T. O. Hodson, "Root mean square error (rmse) or mean absolute error (mae): When to use them or not," *Geoscientific Model Development Discussions*, vol. 2022, pp. 1–10, 2022.

[139] T.-Y. Kuo, P.-C. Su, and C.-M. Tsai, "Improved visual information fidelity based on sensitivity characteristics of digital images," *Journal of visual communication and image representation*, vol. 40, pp. 76–84, 2016.

[140] D. Li, M. Hao, J. Zhang, B. Hu, and Q. Lu, "A universal hypercomplex color image quality index," in *2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings*. IEEE, 2012, pp. 985–990.

[141] M. R. Naufal, C. A. Sari, E. H. Rachmawanto, L. B. Handoko, F. O. Isinkaye, and W. S. T. Al-Dayyeni, "An evaluation of number of pixels change rate (npcr) in symetric cryptography based on data encryption standard (des)," in *2023 International Seminar on Application for Technology of Information and Communication (iSemantic)*. IEEE, 2023, pp. 490–495.

[142] S. A. Güven, E. Şahin, and M. F. Talu, "Image-to-image translation with cnn based perceptual similarity metrics," *Computer Science*, vol. 9, no. 1, pp. 84–98, 2024.

[143] J. Cai, Z. Cao, and L. Zhang, "Learning a single tucker decomposition network for lossy image compression with multiple bits-per-pixel rates," *IEEE Transactions on Image Processing*, vol. 29, pp. 3612–3625, 2020.

[144] A. N. Zemliachenko, S. Abramov, V. V. Lukin, B. Vozel, and K. Chehdi, "Compression ratio prediction in lossy compression of noisy images," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2015, pp. 3497–3500.

[145] S. Kalarikkal Pullayikodi, N. Tarhuni, A. Ahmed, and F. B. Shiginah, "Computationally efficient robust color image watermarking using fast walsh hadamard transform," *Journal of Imaging*, vol. 3, no. 4, p. 46, 2017.

[146] W. Villegas-Ch, J. García-Ortiz, and J. Govea, "A comprehensive approach to image protection in digital environments," *Computers*, vol. 12, no. 8, p. 155, 2023.

[147] A. Kejariwal, S. Gupta, A. Nicolau, N. D. Dutt, and R. Gupta, "Energy efficient watermarking on mobile devices using proxy-based partitioning," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 14, no. 6, pp. 625–636, 2006.

[148] X. Zhao, K. Zhang, Z. Su, S. Vasan, I. Grishchenko, C. Kruegel, G. Vigna, Y.-X. Wang, and L. Li, "Invisible image watermarks are provably removable using generative ai," *Advances in neural information processing systems*, vol. 37, pp. 8643–8672, 2024.

[149] G. Ye, J. Gao, Y. Wang, L. Song, and X. Wei, "Itov: efficiently adapting deep learning-based image watermarking to video watermarking," in *2023 International Conference on Culture-Oriented Science and Technology (CoST)*. IEEE, 2023, pp. 192–197.

[150] Z. Gan, C. Liu, Y. Tang, B. Wang, W. Wang, and X. Zhang, "Genptw: In-generation image watermarking for provenance tracing and tamper localization," *arXiv preprint arXiv:2504.19567*, 2025.

[151] M. Xue, X. Wang, Y. Wu, S. Ni, L. Y. Zhang, Y. Zhang, and W. Liu, "An explainable intellectual property protection method for deep neural networks based on intrinsic features," *IEEE Transactions on Artificial Intelligence*, 2024.

[152] M. Kutter and F. A. Petitcolas, "Fair benchmark for image watermarking systems," in *Security and watermarking of multimedia contents*, vol. 3657. SPIE, 1999, pp. 226–239.