Multimedia Tools and Applications (2024) 83:5531–5591
https://doi.org/10.1007/s11042-023-15806-y

Digital image and video watermarking: methodologies, attacks, applications, and future directions

P. Aberna[1] · **L. Agilandeeswari**[1]

Received: 22 September 2022 / Revised: 22 March 2023 / Accepted: 10 May 2023 /
Published online: 3 June 2023

Abstract
In recent years, internet technology has grown in advance, and multimedia data-sharing growth rates have skyrocketed. As a result, protecting multimedia data in digital networks has become a significant problem. Multimedia data such as audio, text, video, and image are highly used as a data-sharing communication system which demands security, particularly in image and video. Digital watermarking is the one solution that has gained widespread recognition over the past two decades for data embedding in image and video, a key tactic in multimedia tamper detection and recovery. The review tells about the growth rate and data breaches on multimedia data across different applications, which raises the issue of multimedia security. Notably, social network platforms are highly targeted due to their rapid growth, which has created opportunities for data breaches and multimedia manipulation. Here, the forensic field comes into play, where some data-hiding strategies are used to look for evidence of tampering. Even though watermarking techniques can attain security in tamper detection, they face some issues and challenges across various applications. This motivated us to analyze the existing work carried out by data hiding watermarking techniques in the field of multimedia tamper detection in detail and the gap analyzed. Overall, dataset availability, watermarking performance quality metrics, and several image-processing attacks are all explicitly mentioned. This review paper discusses a comprehensive study of the existing system in the field of tamper detection (both in Image and Video) in detail. Also, the development of existing watermarking techniques, issues, and challenges are covered in detail in this paper.

Keywords Social network platform · Tamper detection · Copyright protection · Image watermarking · Image authentication · Video watermarking

✉ L. Agilandeeswari
agila.l@vit.ac.in

[1] School of Information Technology and Engineering (SITE), VIT, Vellore, Tamil Nadu 632014, India

1 Introduction

Digital media (DM) is one of the widely used network platforms. Movies, websites, social media, and online advertising are all examples of digital media. Due to the development of network technology, the growth of digital media data has increased accordingly. In communication channels, data sharing/transfer plays a higher role in the network's system. Even small perturbations added to the image can lead to misjudgment. Among all the multimedia data, image, and video data are primarily used compared to text and audio, which can also be easily tampered with/manipulated using various software like Adobe Photoshop, Picsart, and other editing applications and tools. Security is a top concern in multimedia platforms across multiple applications, including Content Authentication, Proof of Ownership, and Copyright Protection. Securing data from unauthorized users is one of the significant challenges in DM. The attackers target various fields [64] like e-banking, the multimedia industry, social Internet of Things [12], remote sensing [24, 13], social media [69], telemedicine, and e-healthcare [53]. Image manipulation is an uncontrollable issue on the digital platform. Among all the multimedia platforms, social media is one of the highly targeted domains for image manipulation. The contributions of this article are listed as:

1. Digital media is analyzed and pointed out various:
   i. Data breaches in a multimedia platform
   ii. Issues in Social Media platforms lead to high data breaches, which results in the demand for multimedia security, especially Digital Watermarking.
2. Traditional watermarking: Detailed digital image and video watermarking with its properties, datasets, and quality metrics exist.
3. Limitations of conventional and hybrid watermarking approaches for image and video watermarking across various authentication and tamper detection applications are identified.
4. Comparative analysis of various state – of – the – art techniques on evaluation metrics namely PSNR, and SSIM have been presented.
5. Applications: The suitable applications of image and video watermarking namely Content Authentication, Copyright Protection, and Tamper Detection were highlighted.
6. Open Issues and Challenges of the conventional watermarking technique in the real-time scenario are listed with future research directions that might assist new researchers in this domain.

1.1 Real-time application and its issues

The social network serves as a digital platform to communicate and share data with users worldwide. Social network platforms such as Facebook, Twitter, LinkedIn, Instagram, and WhatsApp play a vital role in data sharing. A social network is employed by billions of individuals and has become one of the defining technologies of our time. As shown in Fig. 1, the total number of social media users in 2020 is estimated at 3.96 billion, and the

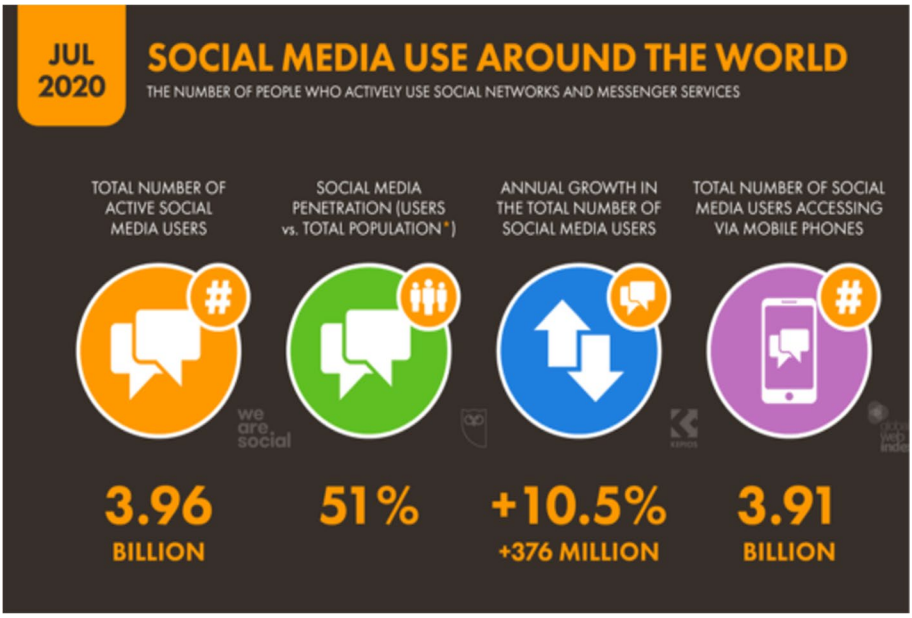


Fig. 1 Usage of social media [123]

annual growth rate of social media increased by 10.5%, which is 376 million [17]. The growth rate is expected to exceed 4 billion users by 2022.

In social media, image/video sharing growth is higher than in other multimedia data. For instance, about 350 million new photos are uploaded on Facebook daily [42]. This small instance shows the drastic growth of image on social network platforms. Because of the unmanageable surge in the usage of digital data, misrepresentation/tampering attack has increased which demand for secure multimedia data (Fig. 2).

Issue 1: The unauthorized data access on social media platform is picturized in Fig. 3, which shows three possibilities: 1) private data migrating to the public

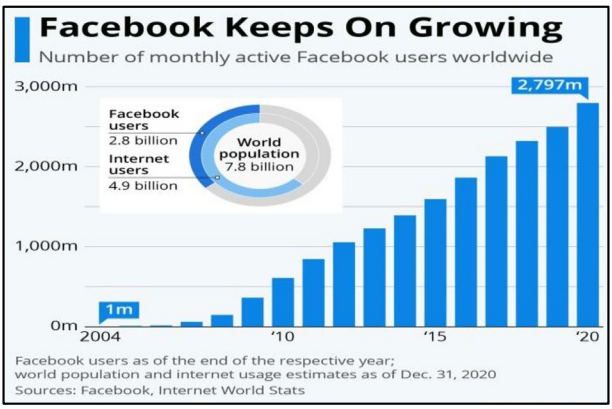


Fig. 2 Growth of Facebook [54]

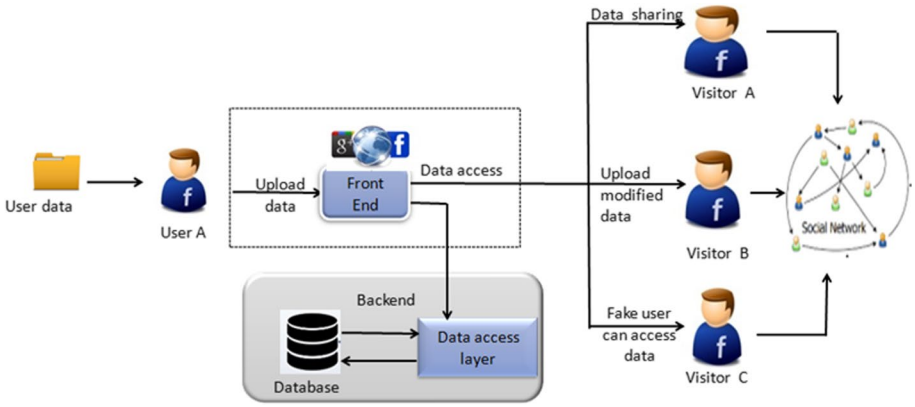


Fig. 3 Shows private data moves to public access in Social Network

environment, 2) need for secure multimedia data, 3) it is evident that reposting (i.e., the images and videos uploaded by Users A on social media platforms are easily acquirable by other unauthorized users) or republishing of data in social media (i.e., if unauthorized user upload multimedia data in their account using 'Tag' option in Facebook, then the uploaded data will be forwarded to other users linked with their account) lead to various problems like malicious attack, content authenticity, or proof of ownership, etc.

Issue 2: Facebook software will make some default changes that user may be unaware of: a) Metadata modification (i.e., if the image uploaded is not in JPEG format, then it changes the image extension to JPEG format to reduce the storage space) [70]; b) original Image id modified by random numbers generation [94]. If the metadata of an image is used as a watermark on social media platforms, particularly on the Facebook platform, then hidden data is unfit to detect tampering.

Issue 3: Auto-bots [161] are a kind of issue in which an individual can create a fake profile through social media software applications. According to Facebook, 6% to 8% of registered accounts are counterfeit accounts [161]. Due to this auto-bots issue, for unauthorized users, social networks became an accessible data accessing platform without the user's acknowledgment compared to all other domains; social media is the highly targeted domain for image misleading.

Social media-related recent data breaches are: On January 20, 2021, a free online photo-editing application 'Pixlr' hacker leaked and stole 1.9 million user records from the database simultaneously [42]. On January 11, 2021, a data breach at 'Socialarks,' a Chinese social media management company, exposed the account details and Personally Identifiable Information (PII) of at least 214 million social media users (including Facebook, Instagram, and LinkedIn), including usernames, profile images, location, Messenger ID, LinkedIn profile link, and linked social media account login, among other things. On different platforms, other data breaches have been recorded [42]. Above listed issue and recent data breach show the serious concerns and the need for secure multimedia data across social media platforms. These serious issues have opened a path for the cybercrime team (i.e., the Forensic field) and researchers to fulfill the demand of users across various applications.

1.1.1 Role of the forensic field

In multimedia data manipulation, forensics plays a crucial role. The forensics team is the branch of image security that focuses on detecting multimedia data manipulation. Typically, the goal of forensic investigators is to [157]: 1) Find whether the image is from the camera or software to verify the image originality and to determine whether the image has been tampered with or not against malicious attack. 2) Locate the tampered region in an image through data-hiding techniques.

Various data-hiding techniques employed by the forensic field to verify authentication and detect image/video distortion are Cryptography, Steganography, and Watermarking. Digital watermarking is defined as the act of concealing information in the original cover image. Unlike cryptography and steganography, confidential information will always stay back in multimedia even after data extraction; WM helps to detect and pinpoint the tampered region across various other applications such as authentication, proof of ownership, copyright protection, and so on. The use of Steganography or Cryptography to transmit multimedia data to the receiver end has significant disadvantages, including the fact that after decryption, everyone can access the data, and it is impossible to prevent data dissemination and illegal copying. As a result, watermarking technique plays a significant role in image security and tamper detection. Watermarking is a known technique because it has existed since the old thirteenth century, but it was implemented in 1992 by Andrew Terkel and Charles in his paper "Electronic Watermark" [101]. Since 1992, Watermark has been slowly extended to various applications like copy control, broadcast monitoring, social media, Healthcare, e-banking, the Military field, etc.

Overall Watermarking technique has three processes: Watermark Generation, Embedding, and Watermark Extraction. Below shown, Fig. 4 describes the working progress of watermarking technique. Original image/video frames are inputted to preprocessing technique to remove noise from the image and to improve the quality of an image. Meanwhile, watermark data *W is* generated (i.e., Metadata, Biometric image, or Feature Extraction) and embedded in the preprocessed original image /video frames, which produces a watermarked image or watermarked video I_W as output. The watermarked data is shared across noisy communication channel 'n,' which can also have a chance for intentional adversary attacks. On the receiver side, they receive noisy watermarked image I_{W_n}. Watermark extraction was performed to detect the adversary attack from noisy data I_{W_n}, which comes under any of these three categories: Blind [88, 121, 129]/ Semi-blind [86]/ non-blind method [38]. No

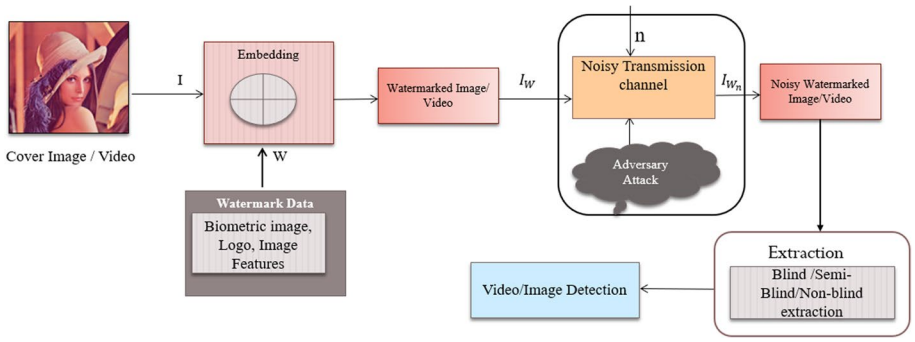


Fig. 4 General Watermarking Process

other information is required to extract the hidden data from the noisy watermarked image I_{W_n}, apart from the secret essential, which is said to be Blind extraction, whereas, in semi-blind, both the original watermark and the secret key are required. All the information like the original image, original watermark data, and the private key are needed, and only then can detect the tamper region; such cases are known as non-blind watermarking techniques.

This section describes existing work carried out by the watermarking techniques in social network platforms. In social network platforms, to enhance tamper detection accuracy and to prove ownership, watermarking techniques generate watermarks using different information for embedding (such as Metadata text information, Biometric data, digital signature, Logo, etc).

Jeffry Bin et al. [32] suggested Metadata information as a secret watermark data implanted in the image and experimented on social media platforms to demonstrate the image's provenance and copyright protection. Metadata such as Date, Time, Manufacturer, Model, Path, Size, etc.… of an image are converted has a binary watermark information. There are two ways to integrate metadata: visible watermarks or undetectable ones. Social network platforms automatically resize and compress the material to conserve storage space. This scaling effect weakens the watermark and tarnishes the embedded data in the uploaded watermarked image on a social network platform. This model's flaw is that compared with other platforms (like Instagram and Twitter), the Facebook platform has a higher chance of the encoded ownership details getting damaged or stripped away from the watermarked image. Digital signature-based watermark data generation was proposed by Gaurav Sharma et al. [32]. Embedding is carried out by changing the least significant bit (3-LSB). Three digital signatures are embedded as watermark data in each image block. The drawback of the system is that it can detect image modification but cannot locate the tampered regions, and it needs more storage space which can degrade the original image quality. Security key-based embedding is suggested in [128] to detect the image against tampering. Watermarked image is shared with the receiver through a social media platform. At the receiver side, using secret key watermark extraction is performed. This method provides good imperceptibility but is not robust against unintentional scaling attacks. Also, acquiring a security key is not difficult for hackers. The result shows that a novel and suitable technique needs to be suggested for real-time application.

The rest of the review article flow is depicted in Fig. 5: Section 2 discusses the watermarking concept, dataset availability for both image and video, quality metrics to measure the performance of the watermark, and various attacks on image and video. Section 3 focuses on applications-based traditional watermarking techniques on robust watermarks and the related work carried out in both image and video watermarking in depth. Section 4 depicts related works carried out in fragile watermarking techniques on both image and video in tamper detection applications, followed by Section 5, which describes other watermarking mechanisms and hybrid domains. Section 6 mentioned existing Deep-Learning based watermarking techniques, followed by Section 7, Open issues and challenges of the watermark. Whereas Section 8 illustrates a comparative analysis. Finally, Section 9 concludes the paper.

1.2 Overall development of watermarking techniques

The evolution of watermarking techniques in the image processing field is shown in Fig. 6, which depicts various watermarking approaches employed for embedding. Visible or invisible metadata embedding, fingerprint images are the primary stage in data hiding. Second

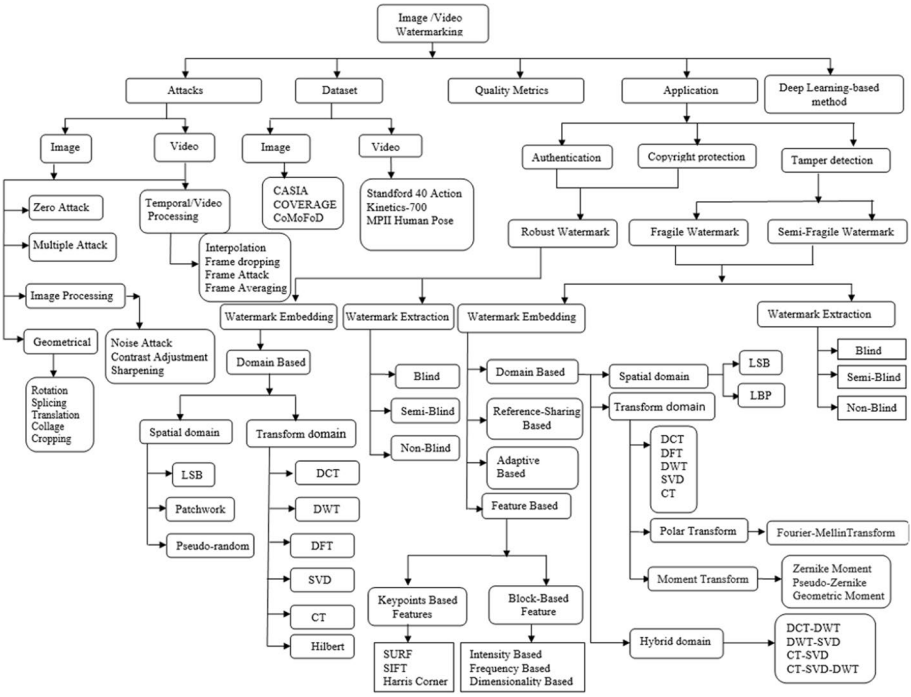


Fig. 5 Taxonomy of Watermarking Techniques

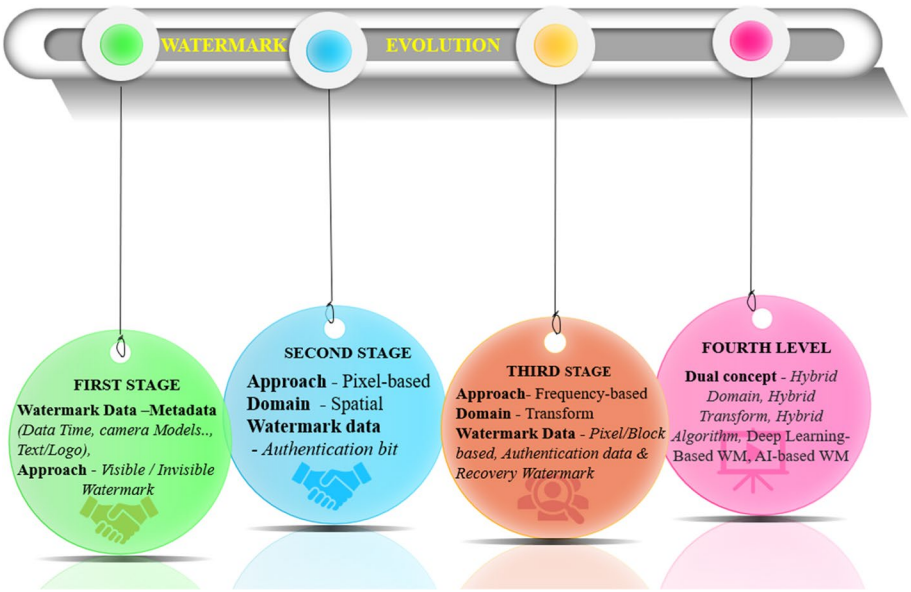


Fig. 6 Overall watermarking concept from the scratch

evolution on the Pixel-based system (i.e., Spatial Domain), where watermark embedding is carried out by various techniques such as least significant bit (LSB) replacement, Correlation-Based, and Patchwork to verify authenticity. Authentication of watermark bits are generated through multiple approaches to verify the integrity of the image. Average intensity computation of the original image sub-block is one instance for the authentication of watermark data 'W'.

In the third stage, they step forward into a frequency-based approach (i.e., Transform domain) to strengthen the image security in case of tamper detection and localization. Watermark data is generated either from spatial values or block-based spatial coefficients or hybrid techniques. Multiple watermarks such as authentication watermark data and recovery data are employed to improve tamper detection accuracy which also increases security levels. The fourth stage of evolution lists the ensemble technique carried out so far. Five different hybrid techniques are 1) Hybrid Domain (Spatial and Transform domain), 2) Hybrid Transform Domains techniques, 3) Hybrid algorithms, 4) Deep learning-based watermark, and 5) AI-Based watermark algorithm which is the most popular trendy edition in the watermarking field.

2 Watermarking concepts

"Digital Watermarking" refers to the process of concealing or embedding data behind an image or video that is invisible to the naked human eye. The Digital Watermarking technique can be widely used for various applications like proof of ownership, authentication, and copyright protection. Digital Watermarking embeds text/images in the original multimedia data without any quality degradation. The embedding information is called watermark data, and the embedding process is called Watermarking technique.

2.1 Types of watermarking

Generally, the watermark can be done in two ways: visible or invisible watermark. A visible watermark is an easy embedding technique, which is perceptible to the naked human eye, but it is not robust against cropping attacks [24]. An invisible watermark is an effective embedding technique that hides the watermark data on the original image, which is imperceptible to human eyes that can resist distortion. For content authentication, invisible watermarking is the most suitable one [43]. Based on this, the watermarking techniques are classified into three categories: Robust watermarking [30, 86], Fragile watermarking [1, 93], and Semi-fragile watermarking [151].

These three types of watermarking are implemented in the Spatial and Transform domains. In the spatial environment [118], the watermark is embedded by directly modifying the host image's spatial coefficient (i.e., pixel values). In the transform domain, the goal is to embed the watermark in the spectral coefficient using various techniques.

According to video watermarking techniques, watermark embedding techniques are classified into three domains: Spatial domain, Frequency domain, and Format-specific (i.e., Moving Picture Expert Group (MPEG) coding- MPEG1, MPEG2, MPEG3, and MPEG4) [68]. MPEG@ is a block-based compression technique that compresses video by predicting its motion. Various watermarking algorithms have been proposed for MPEG2 and MPEG4 coding structures [137]. MPEG4 is mostly used for applications such as video editing and wireless communication system.

2.2 Dataset

The availability of datasets to validate watermarking algorithms against various tampering attacks on Images and Video are listed below. A summary of the dataset is, in short, depicted in Table 1.

Table 1 Description of Image and Video Dataset

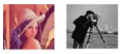
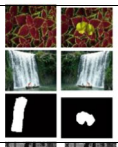
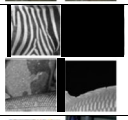
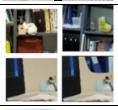
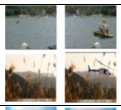
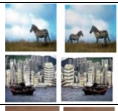
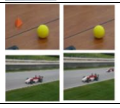
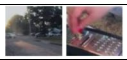
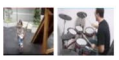
Dataset	Image Size, Total Images, and Videos	Content Description	Sample image and Tampered image
IMAGE DATASET			
Standard test image [22]	256x256, 512x512	Color images and grayscale images	
CASIA-v1.0 [24]	384×256, 800 authentic,921 tampered	JPEG format photos, splicing, ground truth values	
CG-1050(V1) [27]	1050 images, Various dimension	Original and tampered	
Columbia gray [28]	722×480, 933 authentic,912 tampered	Splicing, Copy-paste	
Columbia color [37]	722×480, 183 authentic,180 tampered	Splicing, TIFF format color images	
SIPI [38]	14 are 256x256, 26 are 512x512, 4 are 1024 and 1024	44 images, 16 color and 28 monochromes	
Realistic (Korus) [45]	220 original, 1920×1080	Color image, Tampered image	
CASIA -v2.0 [25]	320×240 to 800×600, 7200 authentic,5123 tampered	Color image splicing detection	
NIST Nimble 16 [76]	500 × 500; 5616 × 3744	Color image, Tampered image	

Table 1 (continued)

BOSSBase v0.93 [77]	722×480, 907 images	Grayscale, PGM format, Cameras images	
CoMoFoD [86]	512×512, 5200 images	Color image, copy-move, JPEG/PNG format	
Grip [100]	768 × 1024 pixels, 3440 images	Copy-paste, Rotation, noise, scaling, JPEG compression	
VIPP [122]	Various size	TIFF uncompressed, JPEG compressed, Splicing	
WildWEB [127]	Various sizes, 10646 images	Cut-paste, copy-move, JPEG/PNG Format	
COVERAGE [132]	400×486, 100 authentic,100 tampered	Copy-move forged images with annotations	
MFC (Media Forensic Challenge) [140]	Various, over 100,000	Splice detection, filtering, and Provenance Graph Building.	
RAISE [142]	Various sizes, 8156 images	Camera native images with high luminance and no compression	
Kinetics-700 **[145]**	650,000	500,000 video clips and covers 600 human action classes	
MPII Human Pose [152]	25,000	410 human activities captured from YouTube videos	
Something-Something V2 (20bn-Something-Something Dataset V2) **[152]**	220,847	168,913 is the training set, 24,777 is the validation set and 27,157 is the test set	
Youtube-8M **[152]**	230K labels	5.6M videos and 3862 classes	
Stanford 40 Actions **[148]**	9532 images	180-300 images per action class	
UCF101(UCF101Human Actions dataset) **[152]**	13,320 video clips	101 categories, 5 class	

2.2.1 Image dataset description

To evaluate the algorithm's performance, a few datasets are used by researchers listed below. The reason behind selecting this dataset for both image and video is: (1) Image-manipulated datasets available publically is attached in the references section with a link (2) Reason for selecting all the below-listed dataset is to provide easy access to all available original and forgery images with different size, various extension, a different image like grayscale, RGB, benchmark gray and color images, low light captured images, all sort of attacked images (3) Also for few datasets ground truth included to validate the performance of forgery detection. Some of them are Image and video datasets listed in the below description:

- **Standard test image (USC-SIPI image database) & SIPI** [122]: are the highly employed dataset distributed in 1977 with various sizes such as 256×256 pixels, 512×512 pixels, or 1024×1024 pixels. All images are 8 bits/pixel for grayscale images and 24 bits/pixel for color images. It is mostly sustained to promote image processing and machine vision research. The standard test image is the widely used benchmark dataset in the image watermarking field, so they compare the performance of the proposed work with the existing work on the same images such as Lena, Baboon, and Barbara **SIPI** [122] also have a few benchmark image sets, which as 44 images with 16 color and 28 monochromes of various size 14 are 256×256, 26 with 512×512, and 4 images with 1024×1024.
- **CASIA-v1.0 dataset** [45] & **CASIA -v2.0** [31]: This is also a highly used dataset, especially in deep learning-based watermarking techniques to train the model. The dataset contains JPEG images of size 384×256 with 800 authentic and 921 tampered images. Whereas CASIA.V2 contains JPEG, BMP, and TIFF format images with 7490 authentic and 5123 spliced of size 240×160 or 900×600.
- **CG-1050 (V1)** [33]: The dataset contains 1050 images of various dimensions. Two directories—training and validation—make up its structure. The original cropped images are kept in one sub-directory, and the altered cropped images are kept in the other. In the cropped image, the manipulated area ranges from 25% to 75%. This dataset can be used to train and test machine learning-based algorithms for categorizing tampered images.
- **Realistic (Korus)** [77]: The dataset contains 220 splicing and copy-move images containing RGB, tampered, and 3-level ground truth maps of dimensions 1920×1080 image. The images were captured by four different cameras: Sony alpha57 (own dataset), Canon 60D (courtesy of dr Bin Li), and Nikon D7000. All images are 1920×1080 px stored in the TIFF format and also include PRNU signatures.
- **NIST Nimble 16** [28]: National Institute of Standards and Technology (NIST) Nimble 2016 included splicing, copy-move, and inpainting tampered and RGB images with dimensions 500×500, 5616×3744.
- **BOSSBase v0.93** [26]: Contains 10,000 grayscale uncompressed images with PGM format, originally designed for research in the steganalysis field with the dimensions of 722×480, 907 images.
- **CoMoFoD** [134]: The 260 forged image sets in the CoMoFoD database of the small dimensions of size 512×512 and large dimensions of 3000×2000 with extension JPEG/PNG format. Translation, rotation, scaling, combination, and distortion are the five groupings divided up based on the manipulation used. All fake and real images are subjected to various post-processing techniques, including JPEG compression, blurring, noise addition, color reduction, etc.
- **Grip** [145]**:** These datasets were chosen because they all include the ground truth images that correspond to each forgery (such as Copy-paste, Rotation, noise, scaling, and JPEG compression), and these images are necessary to assess performance at the pixel level. The 3440 images of size 768×1024 pixels with 80 actual color primary images exported in PNG format. The lowest duplicated region is roughly 4000 pixels, whereas the maximum copied area is 50,000 pixels in size.

- **VIPP Gropu dataset** [142]: 1) **DeepStreets Dataset** – contains road forgery images, **UNISI & UNIFI Dataset –** Has around 700 TIFF uncompressed images, a Universal Attack Against Histogram Based Image Forensics, and so on.
- **WildWEB** [152]: A large collection of forgeries datasets with 10,646 images gathered from different Web and social media sources, including ground truth binary masks, and forged by Cut-paste & copy-move attacks with extensions in JPEG/PNG format.
- **COVERAGE** [132]: dataset forged with copy-move forged images with annotations with dimensions of 400×486, with 200 primary images, of which 100 are authentic, and 100 have tampered.
- **MFC (Media Forensic Challenge)** [85]: Benchmark dataset for digital media forensic challenge evaluations. N datasets were assessed together to progress and deeply analyze the performance of forensic tasks for the past two years. The benchmark dataset contains four major parts: (1) 35 million images and 300,000 video clips from internet sources with labels; (2) up to 176,000 pristine high provenances (HP) images and 11,000 HP videos; (3) approximately 100,000 manipulated images and 4000 manipulated videos from approximately 5000 journal manipulated images and nearly 500 journal manipulated video provenance graph details. (4) evaluation of datasets with ground truth to support challenge tasks in media forensic evaluations.
- **RAISE** [28]: Native images were captured using a Nikon D90 camera with 8170 and high luminance uncompressed images.
- **Columbia gray** [97] **& Columbia color** [98]: The dataset with an image size of 722×480 contains 933 authentic, and 912 tampered (Splicing, Copy-paste) images, whereas, in the Columbia color dataset, images of the same size 722×480 have 183 authentic and 180 tampered (Splicing images) with TIFF format.

2.2.2 Video dataset description

- **Kinetics-700** [76]**:** 700 different human action classes are covered by the 650,000 video clips. Along with human-human interactions like handshakes and hugs, human-object interactions like people playing musical instruments have a minimum of 700 videos in each action class. Each clip is about 10 seconds long and labeled with an action class.
- **MPII Human Pose & Youtube-8 M** [141]: Overall, it has 25,000 video clips with 410 human activities captured from YouTube videos, whereas youtube-8 M has overall 230 K human-verified labeled segments with 1000 classes also contains 5.6 M videos and 3862 classes, each of about 180–300 images per action class.
- **Something-Something V2 (20bn-Something-Something Dataset V2)** [141]**:** Large labeled dataset video clips collection that shows human's basic actions with everyday objects. It contains 220,847 videos, with 168,913 in the training set, 24,777 in the validation set, and 27,157 in the test set with 174 labels.
- **Stanford 40 Actions** [148]**:** Dataset 40 human actions have images with 9532 images in total, with 180–300 images per action class.
- **UCF101(UCF101Human Actions dataset)** [141]**:** The expansion of UCF50 is the UCF101 dataset which has 13,320 video clips that are divided into 101 categories. These 101 categories have five different actions Body motion, Human-human interactions, Human-object interactions, Playing musical instruments, and Sports. These videos run for more than 27 hours in total. A fixed frame rate of 25 FPS with 320×240 resolution of all the videos is downloaded from YouTube.

2.3 Attacks

Multimedia data are manipulated by various attacks 1) Image processing attacks, 2) Geometric attacks, and 3) Multiple attacks. Video-related additional attacks are spatial, temporal, or Video processing attacks (Fig. 7).

Image processing attacks [9] have occurred either intentionally or unintentionally. Various image processing attacks are Noise Attacks (Gaussian noise, Salt and Pepper Noise, Poison Noise), Filtering attacks (Average filter, Median filter & Gaussian filter), Contrast adjustment, and compression.
Noise attack: Noise is added if the image is scaled or compressed intentionally or unintentionally. When it comes to motion-captured images, noise is intentional. Various types of Noise attacks are Gaussian noise occurs due to noise from a natural source. Salt-and-Pepper Noise occurs due to an error in data transmission which looks like black (zero-pixel value) and white (high pixel value) due to distortion in pixel values.

Filtering attack: Filters are applied on the image or video frames that may modify the contrast or blur the image, which affects the image quality, using the filters namely, median filter, and smoothing filter.

1) ***Other Attacks***: The other popular attacks are rotation, scaling, translation, cropping, image sharpening attack, bilinear interpolation, Flipping attack, and collage attack. The geometric attack led to a desynchronization error which means the watermark embedded is still present in the cover image, but the embedded location is changed in the image or video frames.

Rotation attack: The rotation attack will rotate the image either vertically or horizontally or at specific geometrical degrees such as 25°, 50°, and 30°.

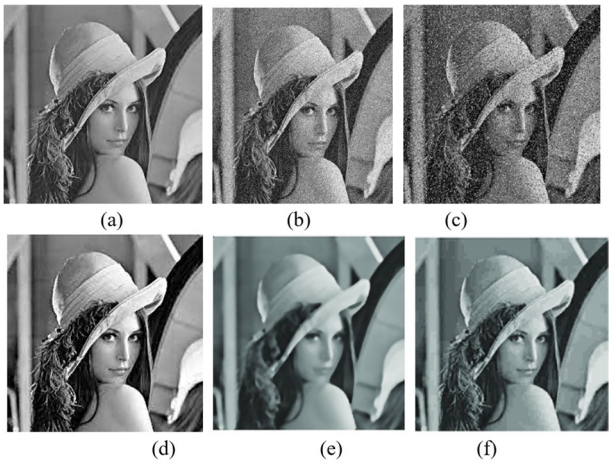
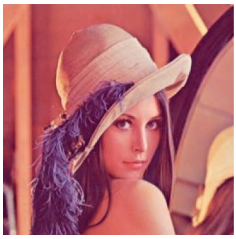


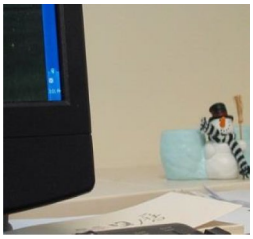
Fig. 7 Lena image (**a**) Original image [127], (**b**) Gaussian noise [72], (**c**) Salt and Pepper Noise [52] (**d**) Contrast Adjustment (**e**) Filtering attack [162] (**f**) JPEG Compression



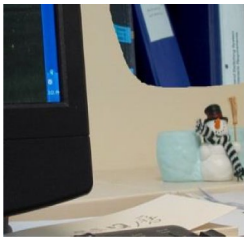
(a) Original Image [127]

(b) Rotated Image(25°)

- **Image Splicing/Cropping attack** will crop the portion of image or region of interest of an image but in the video, the sequence of video frames is cropped and reject all other portions of the video or images.

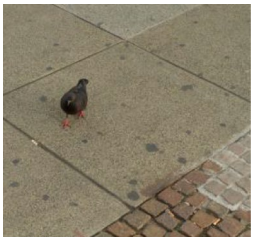


(a) Original Image [37]



(b) Splicing Attack [37]

Copy-move attack: A popular attack in which a part of an image or object is copied from one region and pasted in another region of the same image.



(a) Original Image [38]

(b) Tampered Image [38]

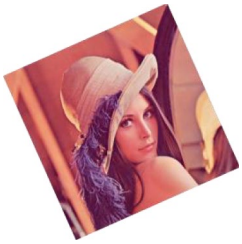
Collage attack: It is a popular image processing tool that allows users to blend two or more images into one source image by default frame available in many image editing software. This attack can be easily performed on the image using photo editing software compared to other forgeries.

Image Sharpening attack: Sharpening improves the definition of an image's edges.

These are some of the intentional attacks as well as highly used manipulation methods because of their ease of use.


(a) Original Image [22] (b) Image Sharpening [22]

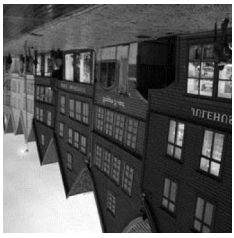
Bilinear interpolation is nothing but resizing the rotated video frames or images to their original size.

(b) Resized Image

a) Rotated Image

Flipping attack: The image is flipped either from left to right or top to bottom.


(a) Original image [38] (b) Vertical Flip (c) Horizontal flip

Scaling attack: A scaling attack is nothing but hiding a smaller resized image in a larger image which will degrade the quality of the image.

2) ***Multiple Attacks:*** The resilience of the embedded image or video is verified by applying multiple attacks such as Copy- paste, contrast adjustment, resizing attack, etc.

a) Original Image

b) Copy-paste, contrast adjustment, Resizing attack in parallel

3) ***Video Watermarking Attack*:** The video-based tampering attacks are classified as Cutting, morphing, and replacement which are instances of spatial tampering attacks. Temporal attacks include frame reordering, frame removal, and frame additions. Tampering attacks can occur at various levels, at each scene, pixels, blocks, frame, and shot [112]. Video processing attacks are termed temporal attacks. Every attack applied to an image also applies to the video, which shows the same result as above, but the only difference is the video frames. Randomly dropping frames from videos is said to be *Frame Dropping,* whereas Frame Swapping is defined as swapping video frames. *Frame Averaging* is nothing but averaging the frames of video.

(a) Noise video attack [7]

(b) Filtering video attack [7]

2.4 Properties of watermark

The watermarking properties are Robustness, Imperceptibility, Payload, and Invertibility.

1. **Imperceptibility**
   After the watermark is embedded in the original image, the watermarked image should resemble the same as the original image. Quality degradation should not be higher. The peak signal-to-noise ratio and Structural similarity index (SSIM) metric measure the similarity between two images to show the embedding effectiveness.
2. **Robustness**
   The ability to withstand compression, noise, and different geometric attacks makes robustness a crucial watermark quality. Using the normalized correlation coefficient (NCC) metric, resilience performance is evaluated.
3. **Payload**
   The watermark capacity indicates the bit-per-pixel data included in the original image. Good robustness will be attained by high payload data embedding, and it also ensures

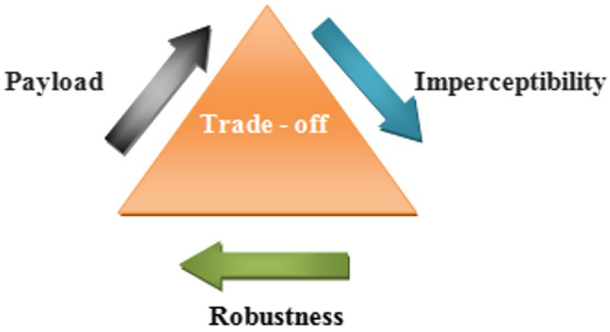


Fig. 8 Trade-off between the above characteristics of the watermark

that the embedding payload should be less than the distortion limit, which won't affect image quality. When payload capacity rises, there will be a conflict between watermarking quality (Imperceptibility) and capacity.

If we embed data in high frequency with less payload capacity, the image gets less degraded but fails with less robustness; if data is embedded at low frequency, it shows high robustness with less imperceptibility. There is always a trade-off between these characteristics, which is considered to be an open issue in the watermark, as shown in Fig. 8.

2.5 Quality metrics

The performance of the watermarking algorithm is evaluated by various metrics, are Peak Signal-to-Noise ratio (PSNR), Mean Squared Error (MSE), Structural Similarity Index (SSIM), Normalized Correlation Coefficient (NCC), Bit Error Rate (BER), Bit Correction Rate (BCR), False Positive Rate (FPR), False Negative Rate (FNR), Learned Perceptual Image Patch Similarity (LPIPS) and so on. The performance will show the efficiency of the algorithm and the quality ratio obtained by the Watermarked image and the Tampered image.

Peak Signal-To-Noise Ratio (PSNR): PSNR is a performance metric representing the quality difference between the original and watermarked images. The PSNR value is expressed in decibels (dB), a value more than 25 to 30 dB is considered as better quality. If the value is below 25 dB, which indicates the performance of the algorithm used to embed the watermark is not acceptable to use. PSNR is computed by taking two Mean Square Error (MSE) images. The PSNR is computed by Eq. 1:

$$PSNR = 10\log10\left(\frac{(\text{Max } gray\ level - 1)^2}{MSE}\right) \tag{1}$$

Mean Squared Error (MSE): The MSE metric defines the error difference between the watermarked image and cover image, whereas PSNR is a measure of cumulative squared error. The mean squared error is the most fundamental way to define PSNR. (*MSE*). *MSE* is defined by Eq. 2:

$$MSE = \sum_{M,N} \frac{\left[\left(I_1(m,n) - I_2\left(m,n\right)\right)^2\right]}{M * N} \tag{2}$$

Where M and N are the numbers of rows and columns of the input images, it squares the difference between two images to highlight errors.

Structural Similarity Index Measure (SSIM): SSIM is a perception-based model for calculating the similarities between two images (i.e., watermarked image and original regarding). The degradation of an image against malicious attack and the similarity between the original and watermarked image is measured by SSIM. SSIM considers the variance, covariance, and correlation between spatial pixels, incorporating luminance and contrast masking. SSIM is considered to be effective if the value is between 0 to 1. Formula to calculate SSIM through Eq. 3:

$$SSIM = \frac{\left(2\mu_x\mu_y + \mathrm{C}_1\right)\left(2\sigma_{\mathrm{xy}+\mathrm{C}_2}\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_1\right)} \tag{3}$$

Where x and y are the original images and the watermarked image, μ_x and μ_y are, respectively, the local means of x and y, σ_x is the variance of x, whereas σ_y is the variance of y, C_1, and C_2 are two variables used to stabilize the division with weak denominator.

Normalized Correlation Coefficient (NCC): To measure the performance of robustness NCC metric is used. It evaluates the similarities between the original and extracted watermarks after distortion. The system is adequate if the value is one or less than 1. Robustness computed using Eq. 4,

$$NCC = \frac{\sum_{i=1}^{n_L}\sum_{j=1}^{n_K}\left(|W(i,j) + W'(i,j)|/2\right)}{n_L \times n_K} \tag{4}$$

Where W and W′ are the binary original and extracted watermark images, and nL and nK are the width and length of the host image, respectively.

Bit Correction Rate (BCR): It's the difference between accurately recovered watermark bits and the total number of bits embedded. If the received message is error-free, then the bit correction rate will be 0; otherwise, it will be close to 1. BCR can be computed using Eq. 5,

$$BCR = \sum_{i=0}^{n} |Oi - Wi| \tag{5}$$

where Oi is the original watermark's intensity, Wi is the extracted watermark's intensity, and n is the total number of embedded watermark bits.

False-negative rate (FNR) The average that the positive region (non-tampered region) predicted as a negative (tampered) value between a non-tampered zone and a tampered region is said to be a false-negative rate (FNR) (i.e., the Tampered region is detected as non-tampered region). Eq. 6 represents as follows:

$$FNR = \frac{No.Of\ pixels\ detected\ as\ non - tampered}{Total\ no.of\ pixels\ detected\ as\ non - tampered} \tag{6}$$

False-positive rate (FPR): The average of identifying the positive region (non-tampered) as a negative (tampered) region is said to be a false-negative rate (FNR) (i.e., the Tampered region seen as a non-tampered region). It is defined by Eq. 7:

$$FPR = \frac{No.of\ pixels\ detected\ as\ tampered}{Total\ no.of\ tampered\ pixels\ detected} \tag{7}$$

Mean Absolute Error (MAE) is another method for calculating the magnitude error average. MAE is similar to MSE, but it measures the absolute value difference between images (i.e., original and watermarked). It's computed using Eq. 8:

$$MAE = \frac{1}{N}\left(\mathrm{O_i} - \mathrm{T_i}\right) \tag{8}$$

Bit error rate (BER): BER is employed to measure the number of bit errors that occurred during watermark extraction divided by the total number of bits embedded, as shown in Eq. 9,

$$BER = \frac{N_{Err}}{N_{Bit}} = 100\left(\frac{Number\ of\ bit\ error}{Total\ no.of\ bit\ embedded\ watermark}\right) \tag{9}$$

Bit Per Pixel (BPP): BPP computes data payload capacity by multiplying the row and column of the image along with no. of bits per pixel of the original image. The number of bits per pixel is computed as given below in Eq. 10.

$$Size\ ofan\ image = rows * cols * bpp \tag{10}$$

Learned Perceptual Image Patch Similarity (LPIPS) [27]**:** The LPIPS metric measures the distance between image patches and the perceptual loss between two images. The goal of LPIPS is to learn more about human perception by increasing the value of SSIM and PSNR to 17. The model extensively uses traditional distortions (such as noise patterns, filtering, and spatial warping operations) and CNN algorithm-based outputs. Because it uses a larger dataset and can use the outputs of real algorithms, LPIPS outperforms other datasets of this type. LPIPS employs a scoring system; a low score indicates high similarity between the compared images, while a high score indicates a significant difference between the images. LPIPS is defined using Eq. 11:

$$Score\ value = lpips\ (img1, img2) \tag{11}$$

Frechet Inception Distance (FID) [27] is a quality metric for determining the quality of a synthetic image by testing the quality and performance of images generated by GAN structures. FID computes two distributions for all activations, both real and generated. Primarily synthetic images are compared to real images and scored the similarity. The generated images are not compared with the original images but rather compared with synthetic images. A low FID score indicates a better image, while a high FID score indicates a poorer image. The FID score is then calculated by Eq. 12:

$$d^2(m, C)\left(m_w, C_w\right) = \left\|m - m_w\right\|^2 + Tr\left(C + C_w - 2\sqrt{C * C_w}\right) \tag{12}$$

Where m and m_w represent the mean value of an original and generated image, C and C_w are covariances of original and generated images.

3 Applications

3.1 Authentication and copyright protection

Authentication and Copyright Protection is an important criteria for the most important application in multimedia data security to prove and verify the originality of the image. In contrast, copyright protection helps to validate authorized users. The problem in a real-world application is that, without the acknowledgment of the actual user, unauthorized users can easily access multimedia data and manipulate it. To solve the problem, adopted watermarking techniques to achieve authentication and copyright protection. For content authentication and copyright protection applications, robust watermarking is highly suggested due to its characteristics.

3.2 Robust image watermarking

Unless the cover image is purposefully changed, embedded watermarks shouldn't be distorted during data transmission; in such circumstances, the watermarking is regarded as robust. Moreover, robust watermarking can validate the authenticity of the embedded image.

3.2.1 Spatial domain-based robust image watermarking

Based on two domains, a generated watermark is integrated into the original image: Spatial [34, 82, 93] and frequency [18, 61, 65, 90]. In the Spatial domain, embed the watermark in the host image by directly modifying the pixels value, which has less computational time and is easy to embed. The primary pixel-based method is applied in the spatial domain without degrading the host image's quality. Different techniques to embed the watermark in a spatial domain are followed: Least significant bit (LSB) [43, 62, 116], Patchwork [21, 149], Local Binary Pattern (LBP) [34], Correlation-based watermarking, spread spectrum method.

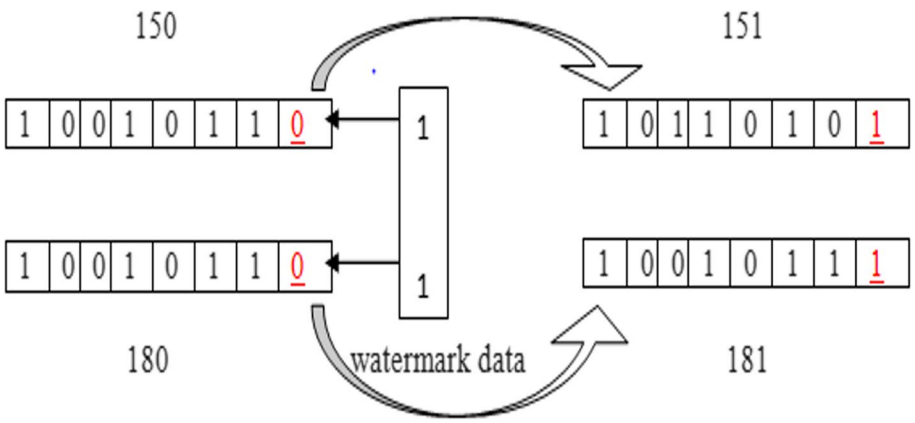


Fig. 9 Process of LSB

In the spatial area of image watermarking, several investigations were carried out. Lamri Laouamer [81] proposed a robust watermarking scheme in the spatial domain to achieve robustness and imperceptibility. Based on the 3×3 subblock threshold value is defined from pixel intensity values to embed the watermark. WM is generated using Weber Law, and pixel-based embedding is carried out. The secret data's payload has been increased to improve tamper detection accuracy, with the help of two watermarks: one for detection and the other for recovery suggested in [30]. Crop-proof watermark is robust for compression, filtering, cropping, and scaling, which is embedded by changing spatial coefficients. In crop-proof watermarks, the JPEG compression algorithm generates the compressed image. A bitmap is employed to embed the watermark. The watermark is detected using bitmap values during extraction, and the tampered image is recovered. For crop-proof watermark extraction, the original watermark image is required for watermark detection.

LSB stands for Least Significant Bit, a steganography method [116] employed earlier than watermarking. LSB is a commonly used simple method in spatial domain techniques that modifies the Least Significant Bit of image pixel values. The maximum embedding limit is 3-LSB because high bit modification will lead to image quality degradation, which results in less imperceptibility. Figure 9 shows an instance to understand how 1-bit LSB works. The watermark data '11' is embedded in the pixel 150 (10010110) and 180 (10010100) by the 1-LSB method; After embedding, the pixel value gets modified to 151(10010111) and 181(10010101). Gil-Je Lee et al. proposed the LSB embedding technique to secure the copyright protection sequence order [82].

Correlation-based Watermarking Technique [68]: Watermark data is generated using Pseudorandom numbers and embedded in pixel luminance value, which is the refined technique of the LSB method. Pseudorandom number embedding is performed based on the host image's supplied "seed" or key value. The gain factor is essential in this algorithm, which decides the quality of the watermarked image. If the gain factor is higher, it degrades the quality and increases the robustness of the host image, or else, increases the quality decrease in robustness.

Patchwork is a statistical process [21] with a Gaussian distribution. Divided into two halves, one is darkened, and the other is brightened. To resist a cropping attack, a pseudorandom binary sequence is generated as a robust watermark embedded in the host image. It is a simple, straightforward method to embed the watermark with less computational complexity. Affine transformation code and feature extraction increased the robustness but less resistance against some geometric attack/distortion. Yeo et al. [149] proposed a Generalized patchwork algorithm. Based on the original image adaptive-patchwork watermark is generated, which attains higher robustness; To improve security, watermark data is scrambled using Arnold transform.

3.2.2 Transform domain-based robust image watermarking

Why Transform domain? The drawback of the spatial domain is that the watermark embedded gets easily tampered with by geometric attacks. The drawback with the LSB technique is that changing the bit value results in a change in color degradation which is not perceived by the human visibility system. Since it has only performed a simple operation, a passive attacker can still easily extract the altered bits [82]. Frequent domain watermarking is introduced to overcome this drawback, where the embedding is done in the frequency band. The image pixel is converted into multi-frequency coefficient bands in the transform domain using various techniques [118]. The transform domain is also called as frequency

domain. To embed the watermark data in the transform domain, we convert the host image to the transform domain using various techniques: Discrete cosine transform (DCT) [62], Discrete Fourier transforms (DFT) [30], Discrete wavelet transform (DWT) [10, 70, 121, 161], Discrete Contourlet Transform (CT) [101, 157] and Singular Value Decomposition (SVD) [55, 133], Quaternion-based DCT/ DFT/ DWT, Hadamard transform, Hilbert Transform [11], Fresnel transform. The advantage of the transform domain is that it is robust for image processing attacks and invariant against geometric distortion.

Discrete Cosine Transform (DCT) Discrete Cosine Transform is one of the important transform domain techniques [36]. One of the characteristics of DCT is that it has high energy compactness. The Discrete Cosine Transform I (x, y) of an image of size M x N is defined using Eq. 13 [87]:

$$\mathrm{I}(\mathrm{x},\mathrm{y}) = \propto(x) \propto(y) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(p,q) \cos\left[\frac{u\pi(2x+1)}{2N}\right] \cos\left[\frac{v\pi(2y+1)}{2N}\right]$$
$$\alpha(\mathrm{x}) = \begin{cases} \sqrt{\frac{1}{N}}, \text{if x} = 0 \\ \sqrt{\frac{2}{N}}, \text{if x} \neq 0 \end{cases}, \alpha(\mathrm{y}) = \begin{cases} \sqrt{\frac{1}{N}}, \text{if y} = 0 \\ \sqrt{\frac{2}{N}}, \text{if y} \neq 0 \end{cases} \tag{13}$$

DCT coefficient bit is computed in the form of sine and cosine functions. On the pre-processed image, we convert the image to DCT transform, and from the DCT coefficient, feature values are extracted as watermark data by Huang. Y et al. [63] mainly designed to resist copy-paste attacks. The low-frequency region of DCT coefficients holds essential features of the original image that are highly concentrated to generate the Watermark (WM) data. The feature matrix is converted to a vector using a zigzag pattern, and feature reduction is made to avoid the overfitting problem. To reduce the feature size truncation method is applied, where the higher frequency vector is truncated. So, feature loss will be less because truncating only the high frequency, as the DCT coefficient low frequency holds important features. The method is sensitive to copy-move forgery, but it fails to detect one or two tampered images in case of a small image size of 16*16, and it fails to be robust against compression attacks. Maheshwari J P et al. [87] proposed a robust watermark in the frequency domain using DCT. Watermark data was generated through the Laplacian pyramid method, which outputs four pyramid images. Out of four Laplacian pyramid images, the first pyramid image size, p0, is reduced to the second image size, p1, through the compression technique. The DCT technique is applied using Eq.13, as shown above. Dhani Ariatmanto et al. [18] proposed DCT based adaptive scaling factor scheme. This method attains high robustness while maintaining the quality of the watermarked image. Arnold transform is applied earlier to the watermark, improving the security before embedding it in the DCT coefficient.

Discrete Fourier Transform: DFT is the process of producing magnitude and phase representation. In mathematics, the discrete Fourier transform is a complex-valued function of frequency in which a sequence of evenly spaced samples is transformed into a discrete-time Fourier transform sequence of the same length and equally spaced samples (DTFT) [68]. The scaling, rotation, and translation are invariant, highly resistant to geometric attack, and can also recover geometric distortion. Though it is highly beneficial, the DFT implementation is higher at a computational cost. Cao, H et al. [30] suggested a robust and reversible color image watermarking algorithm in the spatial domain by fusing Discrete Fourier transform (DFT). From the color image, the channel is selected from which embedding blocks are chosen using a hash function with a secret key. The watermark bit is generated

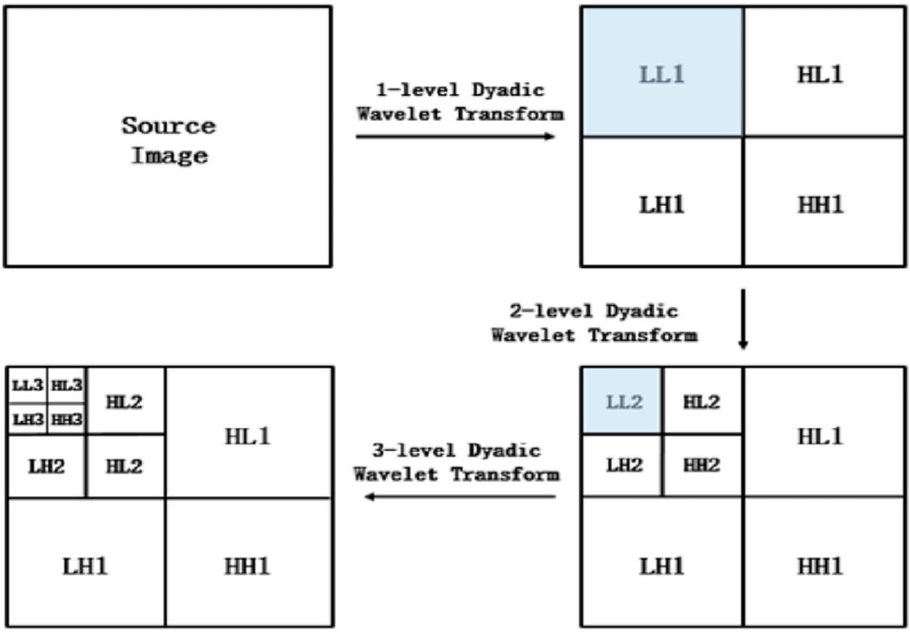


Fig. 10 DWT Process [46]

by encrypting the logo using logistic chaos mapping with a private key and forming a binary sequence. For the selected blocks, the DC component is generated by DFT, and on quantized DC components, the generated watermark bits sequence is embedded 1-bit per block, and the magnitude change of pixels value generates the recovery watermark, which is then implanted to produce a watermarked image.

Discrete wavelet transform (DWT) Discrete Wavelet Transform is a powerful tool in hierarchical decomposition technique. Spatial localization is performed on the image by splitting it into sub-blocks, where the frequency band is divided into four bands. At the first level of hierarchical decomposition, the image is grouped into low, mid, and high-level frequencies represented as LL1, HL1, LH1, and HH1. Among the four bands, LL1 is the low-frequency band, LH and HL are mid-frequency, and HH is the high-frequency band of an image, where the LL1 band contains the actual image output [65]. The second level of decomposition is carried out at the first level of low frequency (i.e., LL1 band), which gets further sub-divided into four sub-bands, which is said to be the second level of DWT; 2level DWT represented as LL2, HL2, LH2, HH2, and it continuous third level decomposition as shown in Fig. 10. Wavelet filters with time-frequency descriptors containing a floating-point coefficient convert the original image to the DWT image. Numerous wavelet filters are utilized, including Haar, Daubechies, biorthogonal, and lifting wavelets. DWT is robust against cropping and rescaling attacks but poor against noise, compression, and geometric distortion attacks. In [65], the proposed DWT watermarking technique for a color image, watermark data embedded in wavelet sub-bands of the blue channel of the RGB that are invariant to various attacks like cropping, scaling, and JPEG compression. The watermark data performance is measured using the normalized correlation coefficient metric.

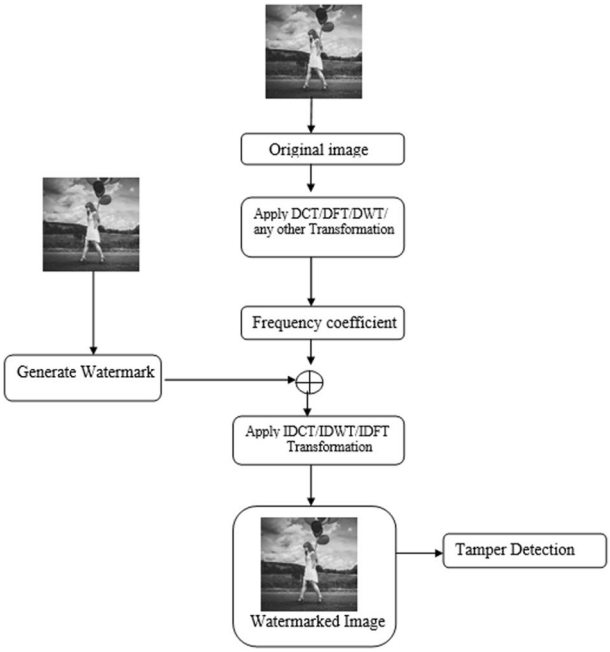


Fig. 11 General Watermarking process in the Frequency domain

The experiment results showed good performance, but the algorithm needs to be improved to reduce the false positive issue.

Alaa Rishek Hoshi [62] presented a robust watermarking technique using a 5-level DWT transform domain approach. Their objective is to improve copyright protection across multimedia networks. The spatial Image is transformed to a wavelet coefficient, and 5-level multi-decomposition is applied. Among all the transforms, DWT shows higher performance due to its signature feature. As shown in Figs. 11 and 5 levels of DWT decomposition are performed on the low-frequency sub-band. Two watermarks are embedded at DWT's second and fifth levels, and inverse DWT is applied to obtain the watermarked image. An extraction process is carried out to detect the tampered region, which needs information about the original watermark and secret key. This kind of extraction process is said to be Semi-blind watermarking.

The general flow of transform domain flow will be in embedding and extracting the watermark, as depicted in Fig. 11. The original image is transformed to a frequency signal using any one of the transform domain techniques, that is, DCT, DWT, SVD, CT, and so on. On the transformed spectral signal, the generated watermark is embedded in the transform domain. Followed by the inverse discrete cosine transform (IDCT) or inverse discrete wavelet transform (IDWT), or inverse Fourier transform (IDFT) applied, which results in the watermarked image. From the watermarked image, we can extract the watermark by just performing the inverse of the embedding process to check and detect the tampering region.

Singular-Value decomposition (SVD) SVD is a highly employed matrix-decomposition technique that can extract the inner feature of the matrix, Where the inner feature is said to be the energy feature of the integer wavelet transform that is highly robust against attack [56]. SVD is a numerical-based intrinsic algebraic tool [97] to decompose the matrix into

Eigenvector and Eigenvalue; that is, SVD is good at the mathematical function where computation is performed in the form of a matrix of any size. It does not limit the matrix size; it can also get further decomposed. By adjusting the value of a matrix's row or column, low perceptible distortion can be created. The image is partitioned into blocks where the SVD of each block is calculated by the below-given Eq. 14:

$$\begin{aligned} \mathrm{X} &= \mathrm{U}\mathrm{S}V^{T} \\ \mathrm{X} &= \begin{pmatrix} u_{1,1} & \dots & u_{1,m} \\ u_{2,1} & \dots & u_{2,m} \\ u_{3,1} & \dots & u_{3,m} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \begin{pmatrix} v_{1,1} & \cdots & v_{1,n} \\ v_{2,1} & \cdots & v_{2,n} \\ v_{3,1} & \cdots & v_{3,n} \end{pmatrix} \end{aligned} \tag{14}$$

Where U and V are two orthogonal matrices, U is the left-singular vector of m, V is +the right-singular vector of n, and S is the singular value matrix [50]. The SVD method achieves high robustness compared with other techniques due to its matrix-based embedding, which is better imperceptible across various malicious attacks, and also it overcomes the drawback of DWT regarding the rescaling attack and reduces false positive problems [56]. Senugul Dogan et al. [44] used SVD based robust watermarking technique for a color image. In biometric applications, data use face image from which iris image is converted to binary LSB significant bit, which is the watermark data, and SVD is applied for face image where the matrix computed and watermark data embedded in S matrix, which increases the robustness. The perception quality is measured using PSNR, which is 62.7 dB, and it is highly robust against geometric transformation especially rescaling attacks; Drawback of DWT is overcome in SVD, but the false positive problem is an issue in the SVD method.

Jing Ming Guo [56] proposed a novel hybrid DWT and Singular value decomposition (SVD) to solve the false positive problem. Watermark data is generated from an image by using SVD and dimensionality reduction principal component analysis (PCA) technique. Host image is divided into block-based embedding employing the spread spectrum concept, which achieves a very good reliable efficiency free from false-positive problems. Compared to all other transforms, SVD is highly efficient, robust against geometric attacks, and controls the false positive rate.

Contourlet Transforms (CT) M.N.Do and M.Vetterli presented Contourlet Transform, which is a unique and efficient multilevel directional decomposition approach and two-dimensional transform method. CT stands for contourlet filter bank, which decomposes directional sub-bands at multi-scale, like DWT, CT directional decomposing technique. The process has two stages shown in Fig. 12: Lowpass multiscale decomposition by

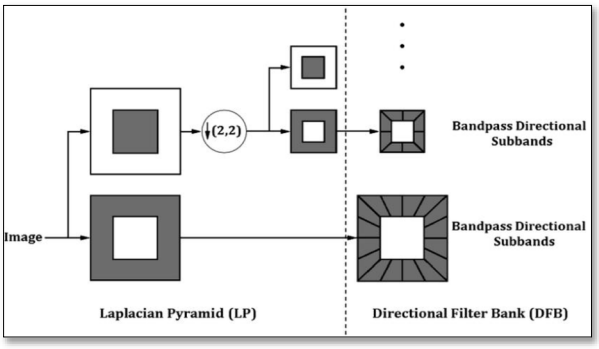


Fig. 12 Contourlet Transform [20]

Laplacian pyramid (LP) and bandpass sub-band is decomposed directionally by directional filter band (DFB), which captures directional information [61]. CT has high characteristics compared to DWT, such as anisotropy and directionality.

3.3 Robust video watermarking

Image Watermarking algorithm and the techniques are extended to video watermarking. Attacks like geometric attacks and image manipulation attack some more additional attacks that occur in video watermarking. Some of the attacks are interpolation, frame switching, frame dropping, frame averaging, and temporal attacks, which include delay and transmission [5]. The watermark embedding in the video is a variant of temporal attacks, due to which it is difficult to attain higher fidelity. The third issue in [118] video watermarking is whether to embed the same watermark in each frame so that when the adversaries try to extract the watermark from different frames, it leads to frames colliding or embedding an independent watermark for each frame, where the attacker would use motionless regions in video frames to remove the watermark by comparing and averaging. Because of the videos' three-dimensional properties, we must consider temporal fluctuation [118].

3.3.1 Spatial domain-based video watermarking

Most of the spread spectrum spatial domain-based work is carried out mostly on copyright protection and robustness applications. Security is a major concern in video watermarking. Image watermarking in spatial domain techniques and methods is extendable to video, where watermark hiding is highly carried out using the Least Significant Bit (LSB) in the spatial domain.

Robust watermarking for video on the spatial domain is proposed by Sharma. V et al. [117] used the LSB technique with a secret sharing mechanism. The generated watermark is embedded through LSB modification, where reference-sharing watermark data is generated using permutation matrix computation. During the extraction process, reference shared image watermark data are regenerated, and by extracting an embedded watermark from LSB, tampered regions are detected in video frames from the watermarked video, which results in high security.

Hartung and Girod et al. [59] proposed spread spectrum watermarking for both compressed and uncompressed video. A watermark is a bit stream that is stretched out, first by replacing the watermark bits. By multiplying the bit stream by a PN sequence and a configurable weight, the final spread spectrum watermark is generated. Following that, a weighted addition is used to embed a watermark in each video frame pixel. Classic spread spectrum technique proposed for MPEG4 standards video frame by Boris Vassaux et al. [137]. Based on the binary sequence (+1) or (−1), the watermark data is embedded in bits, and to improve video watermarking security, Arnold scrambling technique is used along with a secret key.

On compressed videos, Mobasseri et al. [91] proposed a spatial domain video watermarking technique using a direct sequence spread spectrum (DSSS). By utilizing the inherent processing gain of DSSS, it is possible to incorporate the watermark in the raw videos, as well as recover the data from the MPEG decoder (Direct Sequence Spread Spectrum). A single watermark is distributed across various bit planes and scrambled. Watermark

detector uses integration across many planes to utilize this processing gain benefit, which is configurable. A pseudorandom pointer defines the watermark positioning sequence and preempts to which individual bit planes substitute them with a spread watermark in a one-bit plane sequence. Watermark embedded in bit planes of raw video in MPEG and to detect the tampered region decoder decompress and recover the data from MPEG watermarked video. By decompressing MPEG watermarked video, individual frames are extracted by gathering the scattered watermark in bit planes, and it is then matched with the original watermark plane to detect the attacks. Lancini. R et al. [79] presented a spatial robust watermarking scheme for video copyright protection. They used two error-correcting codes, which improved the robustness, and further applied the global mask at the embedding stage to attain better transparency. A binary watermark is embedded on the video frames using pseudorandom sequence by weight parameter *a.*

4 Tamper detection

Tamper detection refers to image manipulation using various software tools without leaving any thumbnails on the cover image. Image manipulation occurs either intentionally or unintentionally. An unauthorized user manipulates the image intentionally, whereas unintentional attacks occur due to the nature of the transmission channel. Various unintentional attack is brightness, contrast adjustment, noise (salt and pepper noise, Gaussian noise), and image blurring [125]. Attacks like Copy-move, rotation, splicing, scaling, and translation attack are some of the geometric attacks which occurs intentionally. For tamper detection applications, a fragile watermark is more suitable than a robust watermark. In a robust watermark, until the original image is secure, the embedded watermark stays robust. When the original image has been tampered with, the embedded watermark will not be robust, and it will get damaged or degraded completely. A fragile watermark is employed highly.to overcome the drawback of a Robust watermark.

4.1 Fragile image watermarking

Fragile watermarking is not strong enough in the textured background against a high tampering rate. Even though the watermarked image is vulnerable to some geometric distortion, unlike the robust watermark, the fragile watermark can detect and locate the tampered area, which is why the researcher highly intended to use this technique. As a result of its sight, it detects the tampered image accurately, and also it recovers the image by extracting the watermark. Usually, watermark data generation or embedding is performed by either pixel-based [38, 70, 86] or block-based [62, 130, 157] techniques, which are more robust against various image processing, video processing, and geometrical attacks [118]. The research field opened a broad way through the fragile technique to achieve better robustness and tamper localization. The fragile watermark embedding process is carried out in the spatial domain [1, 13] and the Transform domain. Most of the recent methods shown in the literature [43, 88, 93, 131, 153] are based on fragile watermarking because of the advantage that it locates tampered regions and is efficiently used for the authentication process.

4.1.1 Spatial domain-based fragile image watermarking

To overcome the drawback of robust watermarking, researchers proposed a fragile watermarking technique in the spatial domain. Pixel-based (i.e., spatial domain) fragile watermarking method is proposed by Zhang et al. [155] for image tamper detection and localization. The binary watermark is created from the image's texture properties, and to increase security, the Arnold transform is applied twice. Generated watermark is embedded in the SVD matrix resulting in high security and perceptual invisibility.

Rakhmawati L. et al. [107] proposed a block-based spatial domain technique using a fragile watermarking approach. By computing the block intensity value and features from the original image, watermark data (i.e., authentication bits and recovery bits) are generated as watermark 'WM' is embedded by modifying the 3-LSB bit, and to avoid quality degradation checksum value is computed and also used two secret keys. This technique is suited for a low tampering rate; if the tampering rate is higher, it is not possible to recover the data. In [73] proposed a self-embedding fragile watermarking scheme in the spatial domain. Self-embedding is nothing but a watermark generated using the original image. The original image is compressed to generate WM data using the absolute moment block truncation coding (AMBTC) approach, which is good at image compression. To improve the quality of the image Optimal Pixel Adjustment Process (OPAP) approach is employed. Watermark embedding is performed in the least Significant Bit (LSB) using the OPAP method. This method achieves better performance than existing work, but it fails at tampering attacks such as filtering and high cropping region.

So far, in the spatial domain-based fragile watermarking technique, WM is embedded by directly modifying the pixel values. In tamper detection applications, along with the pixel-based approach, a novel block-based approach has been introduced. Cumulating all the existing work carried out on the block-based method has two categories: Block dependent (overlapping block) [1] and Block independent (non-overlapping block) method [121]. In a block-based method, WM data is generated from any of these categories like Neighborhood block (i.e., Block Mapping), block-based feature extraction, and block average intensity value. Block-based [138] feature extractor methods are highly preferred for copy-move forgery detection (CMFD) due to their high detection accuracy, and the drawback is high computational complexity.

Rishi sinhal et al. [121] proposed a blind fragile watermarking technique based on a spatial domain for tamper detection and recovery. WM is generated by computing block intensity value along with the secret key embedded in the 2×4 block. Embedding location is computed using pseudorandom binary sequence method range between 0 and 1 and watermark embedded in 2-LSB replacement. By mapping the neighbor block using a smoothing operation, WM is extracted for detecting tampered regions. This kind of extraction is said to be blind watermarking which achieves high robustness. Even though image recovery and tampering detection is achieved at a higher tampering rate (80%) with an accuracy of 99%. This work faces difficulty in recovering the original images against compression attacks, but this technique fails if the mapped neighbor block is also tampered with (i.e., tampering coincidence problem), then it is quite difficult to recover the data. A fragile watermarking block-based method was suggested by A.abdelhakim et al. [1] for image recovery and authentication data. In the K-means clustering technique, the image is grouped into one of these three different forms: horizontal grouping, vertical grouping, and adjacent grouping. The experimental result showed 50% of recovery with high performance, robustness,

and better recovery quality. However, the recovery rate ought to be improved further, and the drawback is authentication system uses only one bit for every block for tamper detection; in case of a tamper coincidence problem, this technique might get degraded. Bitplane selection is suggested by Kim C et al. [74]. They use Absolute Moment Block Truncation Coding (AMBTC) [75] and Optimal Pixel Adjustment Process (OPAP) based color difference model. Obtain a common luminance bitplane from the RGB image followed by k-means clustering that generates common grayscale bitplanes. On each quantized bitplane, 1-bit per block is compressed by color AMBTC and data hiding is processed by the OPAP model. Attained better compression performance than the existing model, and for future work, the quality performance of color AMBTC can be improved.

J.Molina et al. [93] proposed a three-version watermark scheme to overcome the coincidence problem where the authentication data is generated from the block average intensity value by dividing the block of size 8×8 and to improve self-recovery three version watermarks are embedded in sub-blocks. RGB image is converted to grayscale using the YCbCr method to embed a three-version watermark, which obtains a high compression rate. In the extraction process, the halftoning technique is employed, where the first level of the watermark is extracted from an 8×8 block, and the block is sub-divided into 4×4 (16-bits) to extract the second level of watermark data and further sub-divided into 2×2 blocks to extract the third level of data. This method overcomes the existing issue but with less perception quality. Results show high tamper detection and recovery, but there is a chance of quality degradation and tampered region recovery in the case of tampering coincidence problems.

Rakhmawati L. et al. [107] proposed a block-based spatial domain technique using a fragile watermarking approach. By computing the block intensity value and features from the original image, watermark data (i.e., authentication bits and recovery bits) are generated as watermark 'WM'. WM is embedded by modifying the 3-LSB bit, and the checksum value is computed to avoid quality degradation, and it also uses two secret keys. This technique is suited for a low tampering rate; if the tampering rate is higher, it is not possible to recover the data.

4.1.2 Transform domain-based fragile image watermarking

We have mentioned a few works to show how the original image is converted to a transform domain. And in what way are they embedding the watermark, which is said to be fragile?

Recently novel matrix decomposition technique named SVD, a logistic map, and a fragile watermarking approach have been used to locate tampered regions, suggested by Neena Raj N.R. et al. [99]. The 8-bit watermark is generated from each 2×2 block. The logistic map and SVD are used to permute the six Most Significant Bits (MSBs) of each pixel in the block to generate a watermark which is further encrypted using the logistic map to improve security, followed by 2-LSB bit embedding. The experimental results show accurate tampered detection subjected to copy-paste, content removal, text addition, noise addition, vector quantization, and collage attack.

Haghighi et al. [25] suggested a fragile quad watermarking system to attain high image recovery quality. Four chances of watermark generated to recover the image. Two primary reference data are created using lifting wavelet transform and a genetic

algorithm, while the other two secondary data are created using halftoning. Mostly primary reference data are utilized to recover tampered areas with good quality in case of a low tampering rate. If tampering rates are higher, in such cases, secondary WM data is also utilized to recover tampered images, but the results obtained are not good at image quality. A drawback of this approach is that primary reference data are highly utilized; in such a scenario, secondary reference data remains worthless.

Zero watermarking is suggested by [119] for the detection of manipulated document images. Watermark features are retrieved from each non-overlapping 4×4 block of the lifting wavelet sub-bands and combined with the primary watermark to produce the final watermark bits. Watermark bits are placed in the cover image by the mean threshold value calculated from each block. The original watermark is compared with the watermark generated from the tampered image during extraction to verify the authenticity and tamper detection zone.

Using fuzzy logic and Absolute Moment Block Truncation Coding(AMBTC), Manasi Jana et al. [67] proposed a self-embedding fragile watermarking system based on local image features. The blocks are divided into smooth and complicated sections based on pixel correlation using a similarity matrix. Mean 6-MSBs of each smooth block are extracted as recovery watermarks, whereas AMBTC image is used in complicated blocks since it contains a compressed representation of an image. The recovery watermark is embedded in two LSBs of the corresponding mapped block. The visual quality of the watermarked and restored images is high, according to the results, even at a tampering rate of 50%.

4.2 Fragile video watermarking

As mentioned before in the previous section performance of robust watermarks is less compared to fragile watermarks in the tamper detection field. By selecting the number of frames, generated watermark image or video is embedded in a single frame or in multiple frames, which is said to be watermarked video. The integrity issue in the video is solved by using a fragile watermark.

4.2.1 Spatial domain-based fragile video watermarking

Video watermarking also has three processes [96]: watermark generation, watermark embedding, and extraction process. There are two possibilities when it comes to embedding the watermark into a digital video. The first case involves embedding the same watermark in each movie frame. The second scenario involves embedding a distinct watermark in each frame of the video. The second case is not feasible since the video owner must provide watermark data that spans as many frames as possible. Munir et al. [96] proposed a fragile-based spatial domain watermarking for an uncompressed video to achieve video authentication, where a random bit sequence (i.e., Cross-Coupled Chaotic random Bit Generator (CCCBG)) is used to generate binary watermark data. Watermarks are encrypted using the XOR operation and then embedded in each video frame to ensure security. This algorithm is not suitable for compressed video. However, the experiment result shows that it resists cropping, noise, copy-paste, and contrast adjustment, which attain less robustness but are not tired of temporal attacks and video processing attacks.

4.2.2 Transform domain-based fragile video watermarking

Discrete Cosine Transform DCT coefficient-based spread spectrum method is proposed by Nasrollah Moghaddam Charkari et al. [37]. Binary watermark data and the original image are separated into equal-sized blocks. The watermark data block message is placed using the spread spectrum method in the original image of DCT coefficient blocks. When performing DCT conversion, the Hadamard matrix stores the coefficient value. The PSNR metric is used to evaluate the method's performance which attained 51.70 dB. In [102] presented a novel technique to achieve authentication for original raw video. This system used two watermark data, one was to detect tampering, and the second was for tampering localization. The digital signature of each frame computes the hash value in the frequency domain, which generates the primary watermark. The second watermark is generated based on frame numbers and micro-block numbers (i.e., intra-frame and inter-frame). The first seven frames are selected and divided into 4×4 blocks where a binary watermark is embedded in the LSB of the Lempel-Ziv-Welch (LNZ) coefficient. From the tampered watermarked video, the primary watermark is extracted from the quantized DCT signals, but the issue is that the secondary watermark might be degraded. The drawback of this system is the watermark embedded in high frequency, which is not possible to extract from compressed video. To overcome the problem, we can suggest mid-frequency embedding because low frequency degrades the video quality.

Curvelet transform In [9], a novel quaternion Curvelet transform method is suggested for robust and secure video watermarking. The Harris corner detection algorithm extracts the local feature point from QCT coefficients. And to attain a balanced trade-off, an optimized location is computed by the cuckoo search optimization algorithm. This technique is invariant against the geometric attack and results in a good performance with a PSNR value of 69.29 dB.

Wavelet transform In [118] presented a SESAME algorithm with Discrete Wavelet Transform for color video watermarking. Once Successive estimation of statistical measure (SESAME) detects scene changes, it undergoes grayscale image conversion from RGB image. Before embedding the watermark, the 1-level decomposition technique (DWT) was applied to each frame, and the watermark was embedded at two different sub-bands (LL and LH) separately by the blending technique. The performance metric is measured by extracting the watermark from two methods and comparing the LL and LH band watermarks. The comparison result shows the method attains good robustness at both the LL and LH band.

In [6], a new Hilbert transform is used to embed the grayscale watermark image in video frames 2-level Integer Wavelet Transform. An authentication bit is generated from the digital signature to verify the integrity. Before embedding the data, preprocessing is performed by dimensionality reduction technique. This proposed work attains all requirements of a watermark and, after validating the integrity of the image, extracts the watermark to detect the tampered region. The watermarked video attains a PSNR value of 48 dB without any video quality degradation. Himanshu Agarwal et al. [2] suggested a high payload technique for color video watermarking, where the suitable embedding area is computed by key Feature points. Before finding the embedding point, 2D-LWT was applied to the luminance component of each video frame. Five significant interesting feature points are selected by SURF based on the intensity level. To further reduce the computation time, the symmetry

point of the circle method is used by adopting the five interesting key feature points as the center of the circle. In the symmetry of the circle, by considering the quarter portion (i.e., 90° to 45°) using the Pythagorean theorem and Cartesian coordinates on the center point, three more symmetry points are computed. Secret data is embedded at those computed points. For each frame, 20 key points are suggested for embedding the secret data (i.e., 6000 watermark bits are embedded in 300 frames), which attain a high payload. The drawback of this method is that though it embeds high payload data and achieves better imperceptible, it degrades the robustness.

5 Other watermarking mechanisms

5.1 Semi-fragile image watermarking

In a real-time application, when the images are transmitted through the communication channel, unintentional noise and JPEG compression occurs that may affect images; in such cases, the watermarking system will detect the image has a tampered image. To overcome these issues, robust and fragile combined to produce semi-fragile watermarking. Two different semi-fragile are: Recoverable semi-fragile algorithm and Unrecoverable semi-fragile algorithm. The recoverable semi-fragile algorithm can recover the tampered area, whereas an unrecoverable semi-fragile algorithm can only detect and locate the tampering attack [151]. Least Significant Bit modification or by quality factor (QF) approach, Quantization Index Modulation (QIM) approach [78, 15, 105, 154], weight matching functions [144], and error-code [83] are some of the methods to attain a resilient watermark in the case of semi-fragile watermarking.

Semi-fragile watermark is proposed by [16] for securing color images against tampering. Primarily they applied 1 L-DWT on the cover image, followed by a layer selector unit to embed the watermark as they experimented on RGB images. Selected layers are combiner and performed 2 L-DWT. Watermark generated by converting the cover image to YCbCr color space. 1 L-DWT was applied on the Y component, then 8×8 block portioning was performed on the LL sub-band to generate a quantized sub-band. The quantized watermark is embedded in the HH sub-band of selected layers using calibrating factor. This result shows invariant performance against JPEG compression.

Error control code watermark embedding method is proposed in [83] to detect and locate tampered areas. Apply multi-decomposition DWT transform and select HH sub-band to split into non-overlapping blocks. From each block, generate 2D random binary blocks using pseudorandom seed numbers as secret keys. Random binary code watermark is quantized in each block and reconstructed the wavelet coefficient to produce a watermarked image.

A self-recovery technique using a quantization-based wavelet approach is proposed in [105]. From 2-level DWT, recovery WM data is generated from low-frequency coefficient values (i.e., LL2). This LL2 sub-band coefficient is embedded in first-level DWT using Quantization Index Modulation (QIM) [110] approach. The Wavelet coefficient value is rounded to an odd or even quantization level. For the even quantization level, embed "0", and the odd quantization level embeds "1". Only the n most significant bits of each wavelet coefficient from LL2 are used to reduce the payload. The technique is resistant to cropping attacks; however, it is vulnerable to compression assaults. The drawback of this work has

been overcome by Wang X [144]. The Discrete Wavelet Transform technique is suggested for JPEG compression. The watermark data is embedded in the second level of middle frequency. By computing the coefficient energy of 2×2 blocks of LH2 and HL2 sub-bands, the watermark is embedded by matching the energy coefficient relationship. The proposed algorithm is robust against JPEG compression but variant to malicious tampering attacks.

Hazem Munawer et al. [15] put forward an improved quantization-based algorithm in DWT. Embeds the watermark data using a random bit sequence in the second level of the sub-band of DWT (i.e., LL2, LLLH1, LLHL1). The quantization algorithm is not applied to color images and fails to deal with a geometric attack. Another algorithm for the malicious attack is presented by Min-Jen Tsai et al. [135] using the quantization technique. Two-level Wavelet transform is applied to the host image. Quantization technique is applied on low-frequencyLL2 sub-band. A binary watermark is generated quantization method applied on the LL band, which is scrambled by a chaotic system called Toral Automorphisms (Toral Transform) along with a security key. The pseudorandom table with the second key is embedded in the wavelet transform coefficient.

To improve the resilience of watermarking techniques, Xuanping Zhang et al. [154] proposed a new algorithm based on parity quantization Haar wavelet transform. The original image is divided into 4×4 non-overlapping blocks, and the Haar wavelet transform is applied to calculate the Weighted Mean of each block. The initial watermark is generated by a pseudorandom binary sequence using the *Mersenne Twister* algorithm. The maximum low frequency of each block, that is, the max gray spatial value of each block pixel is compared with the original image Weighted Mean to embed the watermark. Watermark detection attains high robustness against lossy compression. In [78], Quantization Index Modulation has been applied to generate security key and watermark data generated by applying the SVD technique on LL bands of RDWT. Hybrid domain RDWT-SVD-based watermark embedded in low frequency, which attained PSNR of 40.59 dB and robust against JPEG compression.

5.2 Reference – Sharing watermarking mechanism

Only if the authorized user shares the clue or seed value or secret key to the receiver end, it is possible to regenerate the original watermark and can detect the tampered region. This is called a Reference sharing mechanism. The advantage of this method is that if the watermarked image is transmitted across the channel and the unauthorized user steals the watermarked data and embeds his/her watermark image, then the actual user can prove his/her ownership by original WM and detect the embedded watermark.

To address the fault tolerance issue, Kang H. et al. [71] suggested a revolutionary secret message-sharing mechanism. They initially create a secret message from the original image's shadow and divide it into n parts based on the number of users. The Direct Bitmap Substitution (DBS) approach is used to integrate partitioned sub-secret messages into the AMBTC compressed cover image. The watermarked image is then uploaded to multiple servers, enhancing the security of secret messages by preventing malevolent attackers from simultaneously getting all of the shadow images. The extraction process should download at least three shadow images from various servers, then uses the Shamir Threshold Method to extract the subsequent messages from the secret images and restore the secret images. For the extraction process, at least three shadow images must be downloaded from various servers. After that, using the Shamir threshold method, the sub-secret messages of the secret images are extracted, and the secret message is then restored using the Shamir

concept of the secret reconstruction algorithm. The benefit of this strategy is that the secret message can still be recovered even if a small number of secret photos are broken during storage or transmission because at least n other users have access to secret images. Also, to strengthen the security of the watermark, a minimum number of users must aggregate their sub-secret messages during the extraction process.

Sharing-based fragile watermarking was proposed by Yu-Jie Chang et al. [35] to achieve authenticity and self-recovery from the tampered image. Vector quantization codebook is used to extract the watermark; the codebook is shared with the receiver based on the threshold value to regenerate the watermark. The index file is created to generate a recovery watermark, and a hash value is computed to generate authentication data, which is concatenated to form watermark data embedded by modifying LSB. Even if the shared data are lost, by using the index value, it is still recoverable because the actual user has reference data, which increases the recovery rate higher.

Fang Cao et al. [29] presented a hierarchical recovery based on a self-embedding reference-sharing mechanism. The cover image is divided into blocks, the LSB method is applied to embed the authentication bit, and the extension ratio of MSB attains high efficiency in the reference sharing bit. To detect various tampering attacks, they modified various extension ratio values and received different imperceptibility result, which is measured using the PSNR metric. This results in attaining high visual quality even with large tampering rates. The advantage of the Reference sharing mechanism is that actual users will have a set of original data with them to prove their ownership.

5.3 Geometric invariant watermarking

So far, we have made a detailed description of various types of watermarking techniques for WM generation and embedding algorithms. Throughout the review, we noticed that attaining invariant watermarks against geometric attacks such as rotation, scaling, and translation is the major challenging problem. Due to the geometric manipulation and synchronization, data loss might occur [143]. Compared with common image processing attacks like noise, JPEG compression, and cropping attacks, generating an invariant watermark against geometric attacks was difficult.

Novel invariant watermark generation to resist geometric attacks is categorized into three principal classes [143]: invariant transform-based algorithms, feature-based methods, and synchronization-based methods.

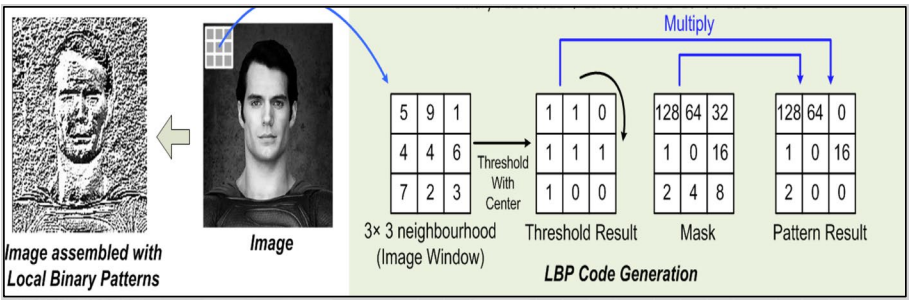


Fig. 13 LBP [100]

The invariant moments such as Pseudo-Zerne Momikents (PZMs) [14], Tchebichef Moments (TMs), Complex Moments (CMs), and Polar Harmonic Transforms, are the techniques to achieve robustness against general geometric attacks. Some of the orthogonally invariant moments are Polar Harmonic Transforms (PHTs) [147], i.e., the Polar Complex Exponential Transform (PCET), the Polar Cosine Transform (PCT), and the Polar Sine Transform (PST). To resist the geometric assaults, Xu et al. [147] suggested a Polar Harmonic Transform (PHT) moment-based to extract invariant orthogonal moment features. The generated watermark is incorporated using quantization methods. It demonstrates that the suggested technique is resistant to filtering, noise, compression, rotation, and scaling attacks.

In the transform domain, feature-based extraction techniques are [41]: local invariant keypoint features, texture-based (local binary pattern (LBP), Gabor transform), histogram of oriented gradient (HOG), polar transform (Fourier-Mellin transform (FMT) [84]), moment transform (Geometric [104, 146], Zernike (ZMs) [41], Hermite transform [40], and so on.

Local Binary Pattern (LBP) in the spatial domain is a 2D surface texture [34]. The local binary pattern is used to extract the feature mask to generate WM data. Figure 13 shows the LBP process; from the topmost 3×3 block of the original image, the center value is selected as the threshold, and the built 3×3 kernels are multiplied with the threshold kernel to produce LBP for a particular block. This process is applied to the entire image, which generates an LBP mask to embed. Watermark is generated using the Local Binary Pattern descriptor of the cover image, which is embedded using a linear interpolation method [80].

In [47], a comparative study has been made with the Zernike moment, which showed more resistance against malicious attack and manipulation. Also, it can locate the duplicated region in the image accurately. Hermite transforms for copy-move tampering detection was presented in [40], which is a sort of signal decomposition technique that is also a special case of polynomial transform. The Gaussian derivative function of the Hermite transform is used to exploit the HVS feature. The Hermite transform and normalization technique achieved high performance in terms of robustness.

The local invariant keypoint-based feature extraction is carried out by a few techniques, such as scale-invariant feature transform (SIFT) [41, 125], speeded-up robust feature (SURF), and Harris corner point detection [9]. SURF [125] is an efficient keypoint feature extractor based on SIFT. Features are invariant against geometric attacks. Watermark detection is based on the key point of the cover image, which is computed using the determinant of the Hessian matrix. An intensity-based scale-invariant feature transform (I-SIFT) approach that can precisely detect the embedding areas is proposed by Fang et al. [48]. Using a lightweight technique, the watermark is repeatedly inserted into various areas. An extraction technique that uses cross-validation and can copy with repeated embedding is used for the extraction process. Experimental results demonstrate the robustness of the suggested approach for screen-shooting. However, feature-based algorithms always have a lower watermark capacity constrained by the feature point and a higher computational cost for processing diverse features.

5.4 Hybrid domain image watermarking

The performance of single domain, i.e., spatial and transform domain approaches, achieve better results on invisibility and resilience. Many approaches based on hybrid domains have been presented to achieve a sufficient performance for watermarking applications.

Hybrid transform domains such as DWT-CT, DWT-SVD [3], DFT-DCT, SVD-CT, and so on [22, 57].

The blind image watermarking hybrid technique is suggested by Mohamed Hamidi et al. [57]. This hybrid technique blends the discrete Fourier transform (DFT) with the discrete cosine transform (DCT). The DFT coefficient magnitudes are used to fulfill the imperceptibility criteria, and DCT is applied to the DFT coefficient magnitudes to fulfill the robustness enhancement requirement. Using a secret key, the mid-band DCT coefficients are modified to embed the watermark, where Arnold transform is used to improve watermark security.

A well-liked hybrid technique called DWT-SVD was introduced by Zhou X et al. [158]. After two-level (2 L) DWT has been applied to the host image, the watermark is incorporated into the singular values of the 2-level DWT sub-band (HL/LH) in this case. In [90], proposed a hybrid-based CT-SVD transform domain which suggested four adaptive watermarking algorithms to attain better imperceptibility and payload To gather singular values based on the direction during partition, the host image has lowpass and bandpass using Laplacian pyramidal transform, as shown in Fig. 9. DWT is used for singular value decomposition, followed by human visual masking and inverse CT to produce watermarked data. The suggested approaches would use an adaptive contourlet technique, which will increase robustness and capacity based on the cover picture size and payload capacity.

A novel hybrid is SVD based support vector machine (SVM) presented by [156]. SVD is used to create the singular value of the original image, watermarks are embedded in integer wavelet coefficients. The experimental finding indicates improved precision against the noisy attack and resistance to geometric attacks. A support vector machine (SVM) and a lifting wavelet transform domain are used to increase the accuracy of tamper detection [139].

The original image is transformed using a three-level LWT algorithm, and the binary watermark is added using an SVM strategy to increase correlation. The randomized coefficient value and other features, such as block-based features, are used to construct the binary watermark. The principal component analysis is used to prevent the problem of overfitting. Compared to the earlier SVD-SVM approach, this work demonstrated a considerable improvement.

The region-based watermarking concept is proposed for normal images, whereas the region-adaptive watermarking technique is suggested by Chunlin Song et al. [124] using the Markov Random Fields algorithm to ensure robustness against various attacks and to improve the imperceptibility. Quad-Tree partition is applied to segment the region of interest and non-region of interest. Watermark generated is embedded in the 2D-DWT transform of a region of interest segment. To make detection more efficient, SVD is applied on DWT sub-bands which results in high robustness.

Highly utilized hybrid transform DWT-SVD method presented by Gao H et al. [51]. The watermark principal component values are generated from the singular value of the scrambled Arnold watermark image. To embed the watermark, DWT is applied on the cover image, and the LL sub-band is partitioned into blocks. The generated principal component is embedded in the singular matrix of each block using an optimal scaling factor which is optimized by an improved artificial bee colony algorithm. During the extraction process, SURF keypoint features of both watermarked and the attacked image are matched. Using the RANSAC algorithm, the tampered wrong key features are detected, and an inverse embedding process is performed to get back the original image.

Table 2 Summary of Image Watermarking based domain and techniques

Watermarking	Domain	Author	Method	Advantage	Disadvantage	Performance Metrics
Robust watermarking	Robust Pixel-based	Alkajith Jyothika [70]	ROI and RONI	Better robustness achieved	ROI and RONI are not possible in all cases	PSNR = 60.1 dB SSIM = 0.995
	Transform Domain	Alaa Rishek Hoshi [62]	5-Level DWT	High robustness	Algorithm performance needs to be enhanced	PSNR = 70.2 dB
	Spatial domain	Lamri Laouamer [81]	Threshold block intensity WM data	Weber Law attains good recovery of the watermark image	Less Robustness against attack	PSNR = 42 dB
	Transform Domain	Sengul Dogan [44]	SVD	Color image with non-geometric attack attains better imperceptibility and robustness	Not worked for geometric attacks	PSNR = 62.7 dB
	Spatial Domain	Yeo [149]	Generalized Patchwork	Effective against attacks	Not robust against random bend attack	BER = 0.3
	Transform Domain	Dhani Ariatmanto [112]	The adaptive Scaling factor, Arnold Transform, DCT	High robustness	Less Imperceptibility	PSNR = 45 dB
	Transform Domain	Mohammed Hamidi [57]	DFT-DCT, blind watermark	Effective against compression attack	Not robust, further need to test against various attacks	PSNR = 61.28 dB
	Robust- Spatial domain	Mohd Aliff [69]	Metadata, visible and invisible watermark (DCT)	Prove the originality of the image	Not robust due to the default resizing algorithm	PSNR = 11.4 dB
Fragile watermarking	Spatial domain	L. Rakhmawati [107]	Block-based,3LSB	Detection accuracy is higher	Vulnerable to image dimensionality change	PSNR = 37 dB
	Adaptive watermarking	Jia Mo [90]	SVD-CT	High security & robustness	Less Imperceptibility	NCC = 1.0

Table 2 (continued)

Watermarking	Domain	Author	Method	Advantage	Disadvantage	Performance Metrics
	Spatial domain	Rishi Sinhal [121]	Smoothing	Higher accuracy at block size 2×2 & 4×4	Higher block size like 8×8 detection accuracy is lesser due to block coincidence	PSNR = 36 dB
	Transform Domain	Behrouz Bolourian Haghighi [25]	LWT & GA- Halftoning	High invariant against some attacks	Less Imperceptibility at a higher tampering rate	PSNR = 46.2 dB
	Transform domain	Gao. H [51]	DWT-SVD, SURF, RANSAC algorithm, Improved ABC optimization	False positive rate reduced	Robustness reduces when attacking density value increase	NCC = 0.99
	Spatial domain	Cheonshik Kim [73]	AMBTC	High Tamper detection accuracy	Block-Based tampering coincidence problem	PSNR = 40 dB
Semi-fragile Watermarking	Spatial domain	Al-Otum [15]	Quantization-based DWT	Suitable for tamper detection technique	Fails to check for the geometric attack, not applied for color image	–
	Transform Domain	Pascal Lefèvre [83]	DWT, Error locating code(BCH)	Robust against JPEG compression	Limited range in Gaussian Noise, quantization algorithm not been applicable for color image and fails to deal with a geometric attack	PSNR = 40.2 dB
	Transform Domain	Wang. X [144]	DWT	Robust against JPEG compression	A variant of tampering attacks	PSNR = 41.85 dB
	Transform Domain	Radu Ovidiu Preda [105]	DWT-Quantization	Resistant to cropping attacks	Vulnerable to quantization and compression assaults.	PSNR = 42.83 dB

Table 2 (continued)

Watermarking	Domain	Author	Method	Advantage	Disadvantage	Performance Metrics
	Transform Domain	Xuanping Zhang [154]	Quantization- Haar wavelet	Robustness against lossy compression, Reduces false alarm pixel	Not resist tampering	PSNR = 34.31 dB
	Transform Domain	Hossein Kourkchi [78]	Quantization Index Modulation (QIM), RDWT-SVD	high transparency, robustness in image JPEG compression schemes, sufficient fragility to malicious attacks	Resistance against Geometric attack is not tested.	PSNR = 40.59 dB

Table 3 Application-based methods with their image size and image type

Method	Application	Original image size	Watermark image size	Image
[70]	Medical Image	256×256	–	Grayscale
[90]	Copyright protection	512×512	128×128	Grayscale
[133]	Tamper detection	512×512	–	Grayscale
[62]	Copyright protection	512×512	512×512	Grayscale
[25]	Tampering Rate	512×512	–	Color image, watermark-Grayscale
[73]	Copyright protection	512×512	–	Grayscale
[152]	Tampering attack	255×255	85×85	Grayscale
[21]	Biometric Image	896×592	50×38	Color image
[22]	Copyright protection	512×512	50×20	Grayscale
[83]	Image Authentication	512×512	–	Grayscale
[15]	Tampering Rate	24×24, 24×24	4×4	Grayscale
[107]	Tamper detection	512×512	512×512	Grayscale
[154]	Image Authentication	512×512	64×64	Grayscale, watermark-Binary image

5.5 Hybrid domain video watermarking

Content authentication for cover video using DWT-SVD suggested by Agilandeeswari L et al. [4]. Wavelet Transform is applied to the Y component of the color video. For each sub-band, block-based singular value decomposition is applied over the selected sub-block to embed the color watermark, and the authentication data is improved by embedding the fingerprint data of the user.

A hybrid combination of transform domains is proposed in [5] using a 24-bit plane slice. Two-level authentication data is generated, primary WM data is generated by scrambling the bit plane slice, and second data is generated by the covariance matrix, and Eigenvector is used as secondary WM which is together embedded in the mid-frequency of DWT-CT. Scramble data is extracted using a key, and Eigenvector is regenerated and compared with the original WM data to check the integrity of the data. If the eigenvector matches, there is no tampering; else, it has been tampered with. This method is an invariant collision attack. Jie sang et al. [113] presented a hybrid technique using DCT-DWT for high efficiency in video watermarking. On the 64 randomly chosen frames of the Y component, DCT applied on each block of size 32×32 is further divided into 2D-DCT (i.e., four sub-blocks). The first, leftmost 2×2 matrix of each frame is concatenated in the sequence, which forms a DC matrix. Based on the embedding strength, the watermark is embedded in the LL sub-band of the DWT transform. Inverse DWT and DCT are performed, which produce watermarked videos. This technique has been compared to the hybrid DCT-DWT-SVD approach, whose performance is highly efficient (Tables 2 and 3).

6 Deep learning-based watermarking approaches

The techniques discussed so far are known as Traditional Watermarking Techniques. Though traditional systems attained good efficiency and robustness towards various applications like authentication, copyright protection, and security, they can't automatically

Table 4 Summary of video watermarking based on domains

Domain	Author	Method	Advantage	Disadvantage	Performance Metric
Spatial Domain	Lancini. R [79]	Spread spectrum, Global masking	Cropping, bit-rate compression, Resizing shows better result	Not robust against Temporal attacks	PSNR = 38 dB
	Hartung and Girod [59]	Spread spectrum	Robust in JPEG compression	Not robust against geometric attacks	NA
	Mobasseri [91]	Direct Sequence Spread Spectrum on MPEG-4	Survive spatial video watermark in MPEG	Not experimented against Security	NA
Transform Domain	Nasrollah Moghaddam Charkari [37]	DCT, Spread spectrum	Robust against compression, brightness adjustment, and contrast, the addition of noise, sharpening, blurring, filtering	A geometric attack is not considered	PSNR = 51.70 dB
	Rupali D. Patil [102]	DCT, Block number, Digital Signature	Robust against brightness adjustment and contrast, the addition of noise	Watermark in high frequency which is not possible to extract from compressed video	PSNR = 39.43 dB
	Himanshu Agarwal [2]	2D-LWT	High payload and better imperceptibility	Fails to attain robustness & perceptibility degrades slowly for a high distortion rate	MPSNR = 60.20 dB
Hybrid Domain	L. Agilandeeswari [5]	DWT-CT-SVD, Eigen vector, 24 1-bit planes	Trade-off attains better balance Invariant against collusion attack	Trade-off becomes lesser in three cases: Background motion is more and video quality brightness is not good and contrast & domination of blue component are lesser	PSNR = 68.09 dB

Table 4 (continued)

Domain	Author	Method	Advantage	Disadvantage	Performance Metric
	Dolley Shukla [118]	DCT-DWT (LL, LH band)	High resilience	A false positive rate is higher in JPEG compression, Gaussian and salt and pepper noise, Frame swapping	PSNR = 77.48 dB NC = 0.917
	Jie Sang [113]	DCT-DWT	High robustness against filters	Geometric attacks has not experimented	NC = 0.9743

Table 5 AI based Watermarking Approach

Author & Reference	Approach	AI method	Role of AI	Advantage	Disadvantage	Performance measure	Applications
Vaibhav Verma & [140]	DWT-SVD	PSO	Compute embedding factor	Reduce the false positive problem	Not robust against geometric attack	PSNR =68.41 dB	Copyright Protection
Issa [66]	SVD, RDWT	GA, Cuckoo search	Robustness and fidelity are optimized	Improve detection accuracy	LSB-based embedding may increase false negative error	PSNR =36.39 dB	Copyright protection, Proof of ownership
Zhou N.R [159]	DWT, SVD	DPSO	Find embedding factor	Improve security, resilience, and imperceptibility	False Negative Problem increases	PSNR = 39.792 dB	False positive problem
M. Rafigh [106]	DCT	GA	Locate the optimal embedding location	Better resilient, Evolutionary algorithm to satisfy both robustness and imperceptibility	Blocking artifacts occurs	NC = 0.97 PSNR = 49.74 dB	Copyright Protection
S. Bhalerao [23]	Spatial-based LSB	DNN	Compressing host image	High-level compression on color image than JPEG and JPEG2000	Performance is not good	PSNR =31.06 dB	Tamper detection
Rao, Y [108]	Basic high pass filter	Spatial Rich Model-CNN,	Extract feature by pre-trained model	Extracted features are robust	Apart from CASIA dataset performance degrades for another dataset	Accuracy = 98.04%	Image splicing & Copy-move attack
Baoru Han [58]	DFT, Mean-hashing, Hermite chaotic scramble algorithm	CNN-VGG19	Feature extraction	Good security, Robustness	Quality degraded at a high Distortion rate	PSNR = 14.07 dB NCC = 0.81	MRI Medical image

Table 5 (continued)

Author & Reference	Approach	AI method	Role of AI	Advantage	Disadvantage	Performance measure	Applications
Mun, S [95]	WMNet	CNN, Detector network-Backpropagation	Training attack simulation features	Robust is achieved against various attacks	Backpropagation slows the detection process	SSIM =0.9706	Tamper detection
Mehdi Rezaei [109]	DCT compression, Reed-Solomon code	CNN	Extract feature	Good security, imperceptibility	False Positive problem	Precision = 0.9776 PSNR = 44.06 dB	Forensic manipulation
Zhu, J [160]	Spatial substitution, Stochastic gradient descent	HiDDeN	Encode-decode and predict using adversary	High robustness and capacity, distortion loss reduced by SGD	Quality degrades	PSNR = 33.55 dB	Tamper detection

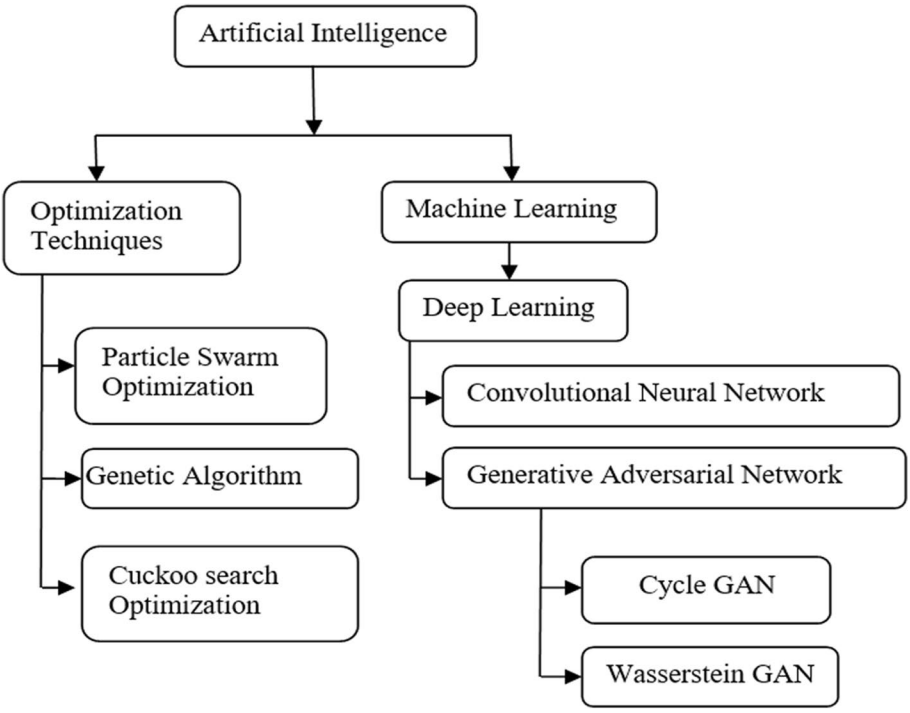


Fig. 14 AI-Based Watermarking Techniques

embed or detect the tampered image, and also, knowledge experts are required to embed mark data. Artificial Intelligence machine-based learning is inspired by human neurons' working system and can do tasks at a human level. Artificial intelligence (AI) has been extended in many fields like natural language processing, image recognition, speech recognition, and Robotics, and it contributed to image security by embedding secret data [19] . Due to this high performance of AI, researchers laid a path to hybridize AI with watermarking techniques to increase digital image detection and recovery accuracy (Table 4).

6.1 AI-based optimization

Artificial Intelligence based optimization techniques plays a role in optimizing the secret data or parameters. The optimization techniques are used for various applications [114, 136] namely, Particle Swarm Optimization Technique (PSO), Genetic Algorithms (GA) [66, 106], Cuckoo Search Algorithm [49, 66], and so on. The role of the AI optimization algorithm in watermarking technique is to: 1) optimize the parameters, 2) optimize the extracted feature values, 3) find an optimized area for embedding the watermark for improving the robustness, and 4) achieve the trade-off between robustness and imperceptibility. Table 5 shows some of the watermarking schemes based on AI. AI-based watermarking optimization and deep-learning techniques are represented in Fig. 14.

A reliable watermarking method employing the Genetic Algorithm optimization methodology was suggested by M. Rafigh et al. [106]. The host image is divided into blocks of size 8*8, and DCT coefficients with a massive difference in frequency values are chosen

for embedding the watermark, which ensures better resilience because selecting high-frequency components in the image will reduce robustness. For each block, a genetic algorithm is assisted in locating the optimal areas, and the fitness value is computed to embed the watermark. The fitness function was used to achieve a balance between these two qualities (i.e., Imperceptibility and robustness).

A DWT-SVD and PSO watermarking scheme was proposed by Verma et al. [140]. Two-level DWT is employed on the host image, and then SVD is used on the mid-frequency sub-bands. For more security, the watermark is divided into two images, each of which is zero-padded to resize it back to that of the original watermark. Using a PSO technique, the segmented watermark pictures are inserted into the host image's singular valued matrix. The appropriate scale factor to divide the watermark image is determined using the PSO. ISVD's two-level IDWT inverse method is employed to generate watermarked images. From the noisy watermarked image reverse process is done.

Issa et al. [66] presented a hybrid watermarking technique using metaheuristic optimization algorithms. The first level of embedding uses SVD transform in combination with GA. Another implementation level uses Redundant discrete wavelet transform (RDWT), and SVD combined with a cuckoo search algorithm. The watermark data is generated by sorting the chaotic map in ascending order, and it permutes the watermark randomly. To embed the watermark into the cover image LSB replacement method is carried out.

Soppari K et al. [126] proposed data embedding by grouping the suitable and unsuitable regions in the cover image using Whale optimization-based FCM clustering technique in a DCT transform domain. DCT transform is applied to the original image by dividing it into blocks, and features are extracted from each block. To select the suitable embedding region, FCM based Least Favorable-based whale optimization algorithm is utilized as centroid, which results in less performance against robustness. The region-based watermarking method is not efficient against malicious attacks because ROI and RONI are not suited to all applications.

Zhou et al. [159] proposed a combination of multiple watermarking techniques, such as lifting wavelet transform (LWT), Discrete Cosine Transformation (DCT), Discrete Fractional Random Transform (DFRNT), and SVD. On the sub-bands of LWT, the DCT transform is applied on each block. To solve inherent problems in the frequency domain (i.e., false positive problem), SVD matrix decomposition is applied to the encrypted image. Meanwhile, the watermark was generated by applying the SVD technique to the original image. Optimal scaling factors are computed using guided dynamic particle swarm optimization (GDPSO) to embed the watermark. The observations result in good in terms of security, resilience, and imperceptibility.

By combining Contourlet Transform, SVD, and PSO, Hatami, E. et al. [60] present a novel hybrid watermarking approach. Both the security and the False Positive Problems are resolved. The host image is partitioned into non-overlapping blocks in the proposed approach, and acceptable blocks are chosen by evaluating the block edge map. The selected blocks that make up the feature matrix that has been chosen are applied to the contourlet transform (CNT). SVD transform is performed on the feature matrix as well as the watermark image. The U and S matrices of the original image are combined to form new singular matrices with a security key. A security key is used to combine the U and S matrices to create new singular matrices. Multi-objective PSO automatically chooses the best scaling factors to enhance robustness, finally, depending on the optimal value of the scaling factor, the watermark image, and the cover image.

6.2 Learning-based approach

Deep Learning (DL) is a subclass of Machine learning, whereas Machine learning is a subcategory of Artificial Intelligence (AI). In the last decade, AI imprints itself in various application areas [111, 136]. Deep Learning (DL) [89] attracted more attention recently, and several outstanding solutions are emerging as a result. Deep learning algorithms span across different field, particularly in the data hiding field, it provides a significant performance which are adaptable and offers a generalized frameworks that lead to the development of various algorithm for image watermarking and steganography problems.

The evolution of DL emerged from Neural networks (i.e., biological neurons vs. Artificial neurons). The basic neural network will have n number of neurons connected to output processing elements, where the input is carried out through weights that move in the forward direction as input to the next layer is known as a Single layer Feedforward network. If an input is interconnected to output layers through hidden layers, then it is a Multilayer network, as shown in Fig. 15. ANN has many types of networks; some of them are Feedforward Neural networks (FNN), Recurrent Neural Networks (RNN), Deep Feed Forward Neural Networks (Long-term short-term memory (LSTM)), and Sensory neural networks (SNN). RNN and LSTM are suitable for Natural language processing (NLP). Below shown, Fig. 13 depicts the Neural network with many hidden layers that are said to be a Deep Neural Network (DNN) or Deep Learning (DL). Deep Autoencoders [95], Convolutional Neural networks (CNN) [39], Deep Neural networks (DNN) [23], BAM Neural networks [7], and Adversarial training (GAN) are some of the Deep Learning algorithms.

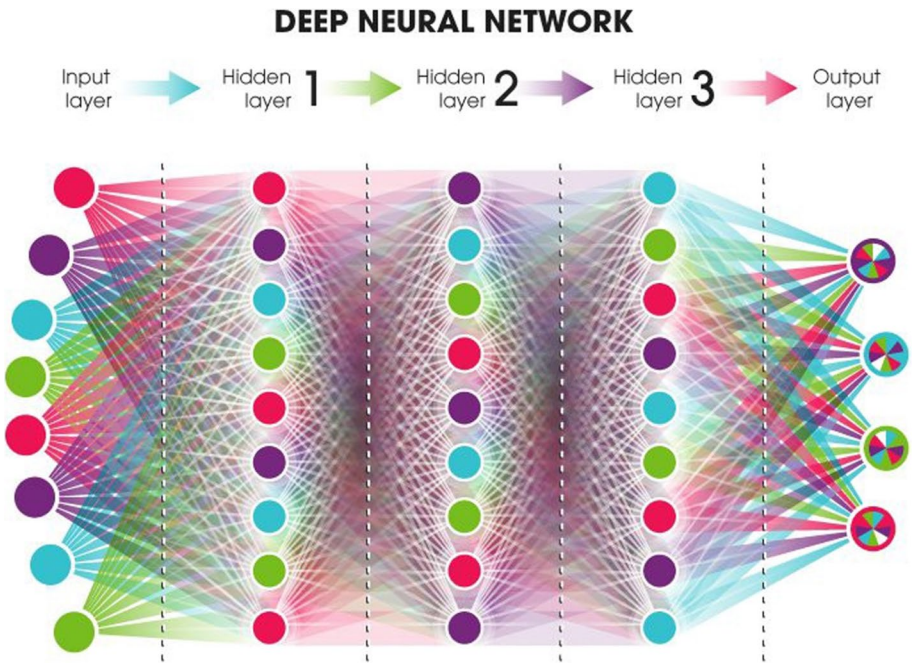


Fig. 15 Architecture of Deep Neural Network [19]

Likely to traditional watermarking, the concept is implemented using a deep neural network (DNN) by training the network [23] with slight modification. In DL models, the usage of CNN is more due to its significant factor that it automatically learns features from the set of image datasets and trains the network model accordingly; Further, it classifies and predicts the test data accurately. Due to its efficiency, it is extended to various applications. In watermarking applications, CNN plays a prominent role in resisting tampered images by training the network with real-world images and a wider range of attacked images, and in testing, it classifies them more accurately than the traditional watermarking algorithm [19]. Surprisingly, Neural Networks (NN) can "detect" information from images that are not perceptible to human eyes. The CNN model's efficiency depends on some assets, such as filters. Based on the filter or kernel, CNN will learn features, and its performance is adjusted accordingly. Apart from these, Bi-directional associative memory networks [8] also play a significant role.

In [23], they used DNN as a compression algorithm instead of JPEG and JPEG2000, which shows great image compression. The compressed DNN image is equally segregated into nine parts, and the watermark image is split into eight parts. Eight-part compressed watermark data is embedded in 9 parts of an original compressed image by selecting the location using a random seed value. The performance is measured by PSNR metric value is about 31.90 dB. Currently, Mehdi Rezaei et al. [109] proposed a DCT compression technique on image blocks and watermark data generated using pre-trained CNN. Usually, the pre-trained VGG19 last layer with 512 kernels extract feature maps, and features are vectorized by a dense layer, whereas here they have used 16 filters at the first layer followed by pooling and dense layer to generate authentication bit codeword. CNN is implemented in two stages: one is on DCT compressed image called ComCNN and the other on the recovery stage called RecCNN. To protect this data from error, the Reed–Solomon error correction code which is a subspace of BCH code, is employed to recover the error from watermark data. Embedding is carried out by the Arnold Scrambling algorithm. The watermark performance is verified by a recall, precision, and PSNR metric, which shows better accuracy.

Robust blind watermarking system designed using CNN model (WMNet) in [95]. WMNet is an image watermarking network used in the CNN model. The progress of watermarking technique is embedding and extraction, whereas the role of CNN is to capture attacked image features to train the network weights to reduce loss function. Block-based embedding is performed by changing the watermark bits where backpropagation and autoencoder are applied. The role of backpropagation is a single detector network, but it causes an error in the computing gradient. Autoencoder is applied further to solve this issue, and the experimental result shows high robustness against attacks.

Instead of a high pass filter median filter is used as a first layer in CNN [39] by Jiansheng Chen et al., which can automatically learn features directly from the image. They were utilizing median filtering as a kernel in the convolution layer of the present system. The problem with CNN is that some remnants are present, making it challenging to identify the tampered area. The first layer of a CNN model in this work was a filter layer that processed the input image and produced its median filtering residual (MFR). Then, the MRF is given as an input to convolutional layers and pooling layers to learn hierarchical representations, and they obtain multiple features for better classification. They tested against cut-paste and JPEG compression attacks, which resulted in good accuracy. Zero-watermarking based convolutional neural network is proposed by Baoru Han et al. [58]. Convolutional Neural networks use pre-trained transfer learning VGG19 to extract deep feature maps from the medical MRI image. The featured image is converted to frequency

by applying the Fourier transform, which computes the mean-perceptual hash value at the primary stage, whereas secondary scrambled logo watermark data is generated using Hermite chaotic neural network. Both primary and secondary are combined with secret keys and embedded in 2D-DFT coefficient values. Followed by the embedding process, results are verified by the extraction process, which is good at security and robustness.

7 Summary

By significantly changing the layers and parameters, so many different CNN models are suggested. Neural networks are prone to varied instances when given an image and a target class: a slight perturbation in a single image pixel will have an impact on the fidelity of the neural network's performance [160]. It may attempt to trick the network into producing inaccurate predictions in the case of some purposeful or accidental attacks, but it should be feasible to extract some useful information from similar altered watermarked photos. It is advised to overcome this problem using the GAN model, which is also a learning-based technique.

7.1 Generative Adversarial Network (GAN)

Apart from the CNN model, researchers mostly adopt the Generative Adversarial Network (GAN) structure. The GAN model contains a generative model and a discriminative model. Discriminator networks in deep data hiding classify the encoded and unaltered images as such. Throughout the learning process, the generative model improves the embedding process and outputs a high-fidelity image, while the discriminative model improves at predicting watermarked images [27]. GAN-based approaches have gained traction, popularized by the HiDDeN model. For data embedding, introduced HiDDeN end-to-end trainable framework [160], which can be applied to both steganography and watermarking. HiDDeN is a new model based on an adversarial instance that has three sectors. An encoder network receives a cover image and marks data, and outputs a watermarked image; a decoder network receives the noisy watermarked image and attempts to reconstruct the watermark data. By injecting noise layers between the embedding and extraction phases and forcing the model to learn to survive noisy picture transmission, the adversary finally determines whether a given image includes encoded data. The stochastic gradient descent technique is applied to reduce the data loss.

Variations in the GAN model lead to two novel models, Wasserstein GAN and CycleGAN. Normally in traditional GAN, the discriminator predicts the probability that the output image is real or fake. But in WGAN, instead of predicting it by discriminator, critic scoring shows the originality of the image. In [103], a Critic score-based blind watermark-based GAN model is designed. The architecture consists of six main components, three are training neural networks called an encoder, decoder, and adversarial critic performed on the YCbCr image, and the other three are additional components called nosier, which is used to perform attacks on the encoded image (watermarked image) and message propagator P and message translator T employed as two deterministic algorithms. Convert the secret message from tuples to binary message and spread the message randomly in the spatial domain, output as an encoded image (Watermarked image). During the training process, attacks are applied, like cropping, JPEG compression, subsampling, and so on. The

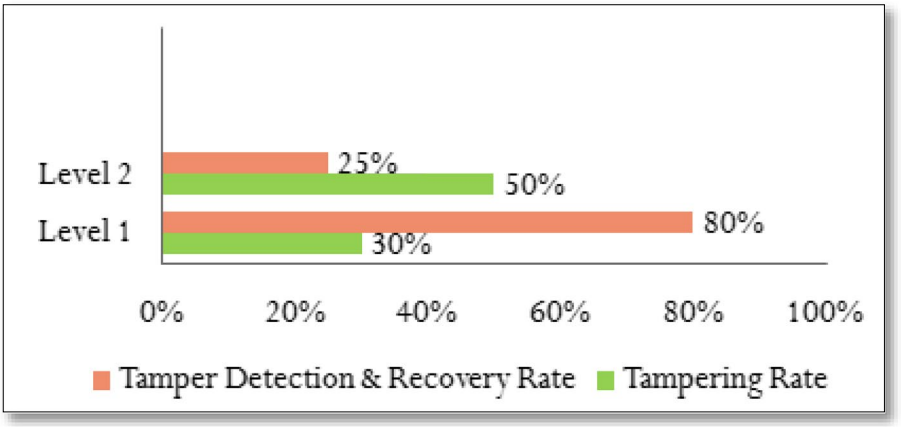


Fig. 16 Tamper detection and recovery accuracy rate

detector extracts the message and computes the critic score to compute the originality of the image, followed by a message translator that extracts the original message from the detected message.

Normally, the GAN model contains one generative-Discriminative phase (encoder-decoder), Small modifications in the GAN model that as two generative and discriminative phases known as CycleGAN. In CycleGAN, an attention-based mechanism is presented by Yu C et al. [150] to find the suitable embedding area on the cover image. The attention mask is generated by the attention generative model and represents the sensitivity of cover image pixels. Instead of hiding binary messages, they embed the secret image. Feeding the attention mask, secret image, and the original image as input to GAN to produce a target watermarked image, which will be similar to the original image. Target image generated by learning mapping function by the generative model, to perform the extraction process, the mapping function is required by the discriminator to extract the secret image which leads to this cycle adversarial training process. During the extraction phase, a secret image is extracted from the secret generative model. The work progress of two discriminator models is performed by computing the cycle adversarial loss function. The first discriminator model (i.e., original image discriminative model) ensures that both the original and the target image are not distinguishable, and at the same time, the extracted secret image (i.e., Watermark image) extracted is fed as an input to the second discriminative model (i.e., secret discriminative model) which check whether both the images are indistinguishable which results in high imperceptibility.

Attention-based tamper detection:

The effectiveness of these traditional techniques falls short in real-world applications such as social media platforms. To handle this multiple manipulations of images, the CNN-based model will perform less on multiple manipulated images. To overcome this issue, they suggested a CNN-based attention mechanism to look into multiple manipulations by Bhuvanesh Singh [120]. Deep Learning CNN learns intrinsic features, and the learned feature vector is given as input to the attention mechanism to train and predict tampered images. The result can detect tampering attacks like text-editing, face-swapping, copy-move, splicing, and mirroring. Local Interpretable Model-agnostic Explanations (LIME) are utilized to locate the manipulated region in a tampered image. The model has achieved

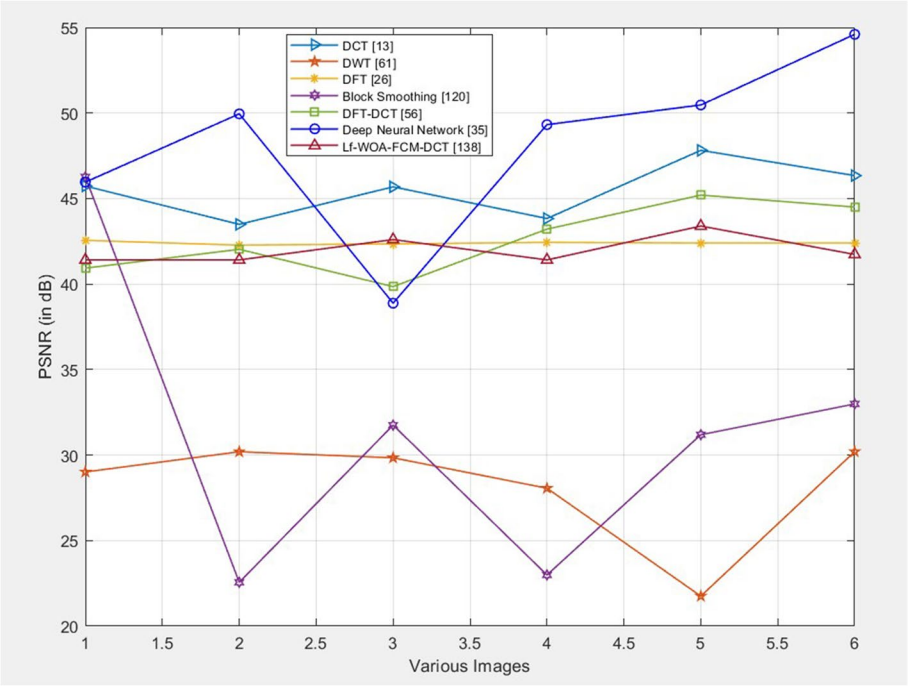


Fig. 17 Various model's performance PSNR Comparison

an accuracy score of 94.7% on the CASIA 2.0 dataset, which is better than the previous CNN works in tamper image detection. They tried this technique on the real-world Twitter dataset and acquired 83.2% over the real world. The experiment proves that the proposed model can accurately detect forged images over social platforms.

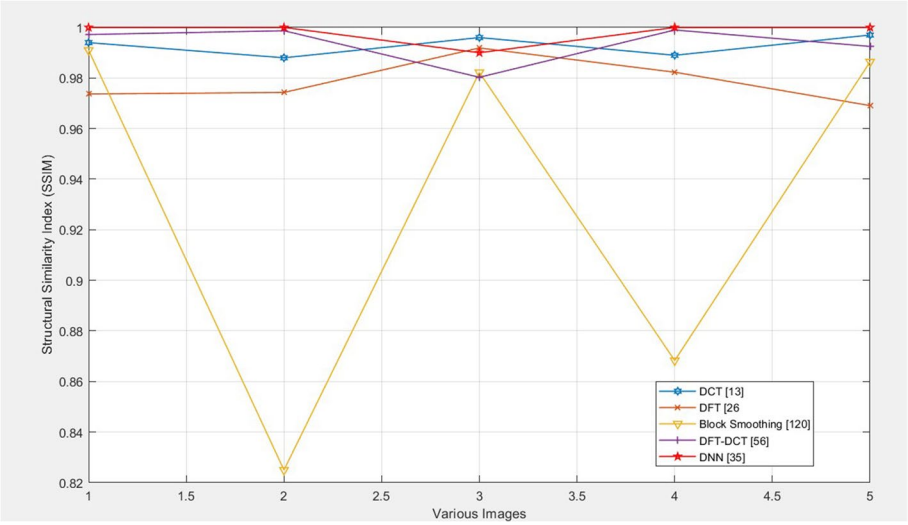


Fig. 18 Various model's performance SSIM Comparison

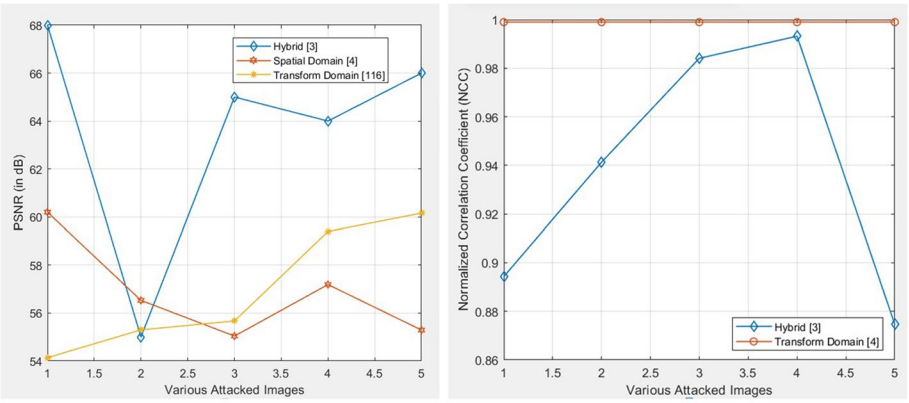


Fig. 19 Various models performance Comparison in terms of (**a**) PSNR & (**b**) NCC

8 Open issues and challenges

This section dictates the open issues and challenges in the watermarking techniques based on the above study,

- The trade-off between the characteristics of the watermark is a major concern.
- If the image tampering rate is higher, then the tamper detection accuracy and image recovery quality will be lesser. In watermarking techniques, many methods were emerged to detect tampered images, but they are incapable of reaching the higher tamper detection accuracy and image recovery quality in case of higher Tampering rates, as shown in Fig. 16.
- Once the watermarked image is uploaded to a social network, the embedded data gets tarnished. So, the robustness and security of watermarked images are still a challenge in the case of the Social Media Data Sharing platform [69].
- The non-blind watermarking scheme will not be suggested for real-time application.
- None of the semi-fragile algorithms meets all the characteristics of watermarking.
- The existing Semi-fragile technique is undistinguishable in case of intentional and unintentional attacks [151].
- Existing semi-fragile technique enhancements continue to focus on content authentication, which is not focusing on tampering attacks, and image recovery [151].
- So far, grayscale or binary image is embedded in the original image. None of the existing systems has used color image watermark data to overcome color image degradation.
- In video watermarking, the challenge is to create a robust algorithm that can withstand compression methods while retaining high perceptual quality [102].
- For a real-time application like a stock photo database, it is difficult to embed the watermark in each image. For such cases, the deep learning-based algorithm can be suggested.
- Both in neural networks and adversarial training, they trained the noisier network with one individual attack; multi-attack is not concentrated so far.
- Even though GAN achieves better results, training through GAN is very hard compared to CNN.

9 Overall comparative analysis

The recent watermarking models (deep learning, hybrid, and conventional methods) from the entire literature review are compared in terms of PSNR, SSIM, and NCC. It is inferred from Fig. 17 on PSNR results are that, the DNN-based watermarking model [23] achieved superior imperceptibility performance than the traditional and block-based smoothing models, followed by the DCT [18] and hybrid DFT-DCT [57] models. In contrast, SSIM results in Fig. 18 reveal that the hybrid DFT-DCT model and DNN-based watermarking model outperformed more traditional models. Finally examined, the performance of video creation based on domain, and from Figs. 19a and 19b, it can be seen that the hybrid model [5] achieves better imperceptibility performance in comparison to the conventional model transform domain [2] and spatial domain [117]. Figure 19b represents fragility, and hybrid both achieve high robustness that is nearly 1, but when the tampering rate value increases, the performance of the hybrid model degrades slightly.

10 Conclusions and future research directions

Securing multimedia data in many applications has become difficult in the digital era. This study examines the use of watermarking techniques in several contexts, including content authentication, copyright protection, and tamper detection. The research field has evolved from the basics to the most advanced stage in watermarking applications. The evolution of watermarks began from metadata-based watermark generation to various methods like adaptive-based, reference-based, feature-based, block-based, hybrid-based techniques, etc. In traditional watermarking, the block-based technique is more effective than the pixel-based (i.e., Spatial domain) technique. Robust watermarking is more efficient in the transform domain than the spatial domain because the performance is limited in the spatial domain. Transform domain-based robust watermarking obtained higher effectiveness not only in tamper detection but also in locating the tampered region; it also resists against few geometric attacks and improves visual quality. A drawback of Robust watermarking is that it is effective against image processing attacks but variant to some geometric attacks. Because of this, the fragile system has opened a broad way in the spatial domain and transform domain. In fragile watermarking, though watermarking is fragile to image manipulation attacks, it is accurate in detecting the tampered region and shows better results at tampered image recovery. In the case of an improved fragile algorithm, the block-based method attains good quality. Apart from all the transform domains techniques reviewed in the fragile-based technique, wavelet transforms especially DWT has good invariant property against geometric attack, but false positive is high. To solve the false positive problem, SVD techniques are employed. A combination of DWT-SVD is highly used in existing systems as it shows better results. In semi-fragile watermarking, JPEG compression and JPEG lossy compression attain high efficiency by quantization techniques. The issue in the existing work in semi-fragile techniques is it is difficult to distinguish between malicious and non-malicious attacks. Semi-fragile is highly looked into content authentication, but proof of ownership and copyright protection is also an urgent requirement in the digital world.

One thing that stands out among all the approaches is the hybrid technique. Instead of a single technique, hybrid techniques perform are more efficient and highly robust. The hybrid domain-based (i.e., spatial and transform domain) watermarking algorithm attained more robustness and high image quality against tampering rather than using individual transform

techniques. However, image security remains a big challenge in attaining ownership and authentication for data-sharing platforms. So far, DL techniques have trained the network with individual attacked images, and we can extend it to train multiple attacks to improve robustness. A hybrid domain-based deep learning algorithm can be used in real-time applications to achieve better security in watermarking techniques. To enhance the performance of video watermarking, ensemble watermarking, and deep learning-based methods are suggested for future research direction. For video data, we need to find a novel robust algorithm to tolerate compression methods, and watermark embedding capacity needs to be increased in the video domain. For future direction, researchers can attempt to use ensemble approaches to achieve proof of ownership and improve watermark security in social networking platforms.

Deep learning techniques are now being researched in a variety of computer vision applications to increase accuracy. Researchers integrated deep learning with watermarking techniques to enhance tamper detection and recovery over traditional watermarking approaches. Convolution neural networks (CNNs) are a crucial part of deep learning and are employed in various applications, such as image classification and object recognition. The disadvantage of watermarking-based deep learning techniques is that even a small change to an image pixel will reduce the neural network fidelity performance. Transformer evolved from the success of the Natural Language Processing (NLP) task in machine translation, which gradually attracted the attention of computer vision applications, much like CNN and the deep learning technique. Deep learning techniques are now being explored in several computer vision applications to obtain more accurate results. But, the use of transformer models is limited to a few computer vision applications, such as Animal Detection and Classification [10], Land-cover classification [115], multispectral satellite images [92], etc. Researchers can even incorporate the transformer model for the watermarking techniques to achieve Copyright Protection and Content Authentication applications like CNN models.

Data availability Data sharing does not apply to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interests/competing interests The authors have no Funding and/or Conflicts of interests/Competing interests.

References

1. Abdelhakim A, Saleh HI, Abdelhakim M (2019) Fragile watermarking for image tamper detection and localization with effective recovery capability using K-means clustering. Multimed Tools Appl 78(22):32523–32563
2. Agarwal H, Husain F (2021) Development of payload capacity enhanced robust video watermarking scheme based on symmetry of circle using lifting wavelet transform and SURF. J Inf Secur Appl 59:102846
3. Agilandeeswari L, Muralibabu K (2013) A robust video watermarking algorithm for content authentication using discrete wavelet transform (DWT) and singular value decomposition (SVD). Int J Sec Appl 7(4):145–158
4. Agilandeeswari L, Muralibabu K (2013) A novel block based video in video watermarking algorithm using discrete wavelet transform and singular value decomposition. Int J of Adv Res Comput Sci Soft Eng 3(4)

5. Agilandeeswari L, Ganesan K (2016) A robust color video watermarking scheme based on hybrid embedding techniques. Multimedia Tools and Applications. 75(14):8745–8780
6. Agilandeeswari L, Ganesan K (2016) An efficient hilbert and integer wavelet transform based video watermarking. J Eng Sci Technol 11(3):327–345
7. Agilandeeswari L, Ganesan K (2016) An adaptive HVS based video watermarking scheme for multiple watermarks using BAM neural networks and fuzzy inference system. Expert Syst Appl 63:412–434
8. Agilandeeswari L, Ganesan K et al (2016) A bi-directional associative memory based multiple image watermarking on cover video. Multimed Tools Appl (Springer) 75(12):7211–7256
9. Agilandeeswari L, Ganesan K (2018) RST invariant robust video watermarking algorithm using quaternion curvelet transform. Multimed Tools Appl 77(19):25431–25474
10. Agilandeeswari, L., & Meena, S. D. (2023). SWIN transformer based contrastive self-supervised learning for animal detection and classification. Multimedia Tools and Applications, 82(7), 10445-10470.
11. Agilandeeswari L, Ganesan K, Muralibabu K (2013) "A side view based video in video watermarking using DWT and Hilbert transform," Security in computing and communications, Communications in Computer and Information Science (CCIS) series – Springer, p. 3
12. Agilandeeswari, L., Paliwal, S., Chandrakar, A., & Prabukumar, M. (2022). A new lightweight conditional privacy preserving authentication and key–agreement protocol in social internet of things for vehicle to smart grid networks. Multimedia Tools and Applications, 81(19), 27683-27710.
13. Agilandeeswari L, Prabukumar M, Radhesyam V, Phaneendra KLB, Farhan A (2022) Crop classification foragricultural applications in hyperspectral remote sensing images. Appl Sci 12(3):1670
14. Agilandeeswari L, Prabukumar M, Alenizi FA (2023) A robust semi-fragile watermarking system using Pseudo-Zernike moments and dual tree complex wavelet transform for social media content authentication. Multimed Tools Appl 1–53
15. Al-Otum HM (2014) Semi-fragile watermarking for grayscale image authentication and tamper detection based on an adjusted expanded-bit multiscale quantization-based technique. J Visual Commun Image Represent 25(5):1064–1081
16. Al-Otum HM, Ellubani AAA (2022) Secure and effective color image tampering detection and self restoration using a dual watermarking approach. Optik 262:169280
17. Appel G, Grewal L, Hadi R, Stephen AT (2020) The future of social media in marketing. J Acad Market Sci 48(1):79–95
18. Ariatmanto D, Ernawan F (2022) Adaptive scaling factors based on the impact of selected DCT coefficients for image watermarking. J King Saud Univ-Comput Inform Sci 34(3):605–614
19. Asiri S (n.d.) " Brief Introduction to Artificial Neural," Meet Artificial Neural Networks. Brief Introduction to Artificial Neural… | by Sidath Asiri | Towards Data Science
20. Azizi S, Mohrekesh M, Samavi S (2013) Hybrid image watermarking using local complexity variations. In: 2013 21st Iranian Conference on Electrical Engineering (ICEE). IEEE, "Contourlet Transform," (n.d.), pp 1-6. https://www.researchgate.net/figure/The-contourlet-transform-consist-of-LP-and-DFB-part_fig2_257547896
21. Begum M, Uddin MS (2020) Digital image watermarking techniques: a review. Information 11(2):110
22. Begum M, Uddin MS (2020) Analysis of digital image watermarking techniques through hybrid methods. Adv Multimed 2020:1–12
23. Bhalerao S, Ansari IA, Kumar A (2021) "Analysis of DNN based image watermarking data generation for self-recovery," 2021 international conference on control, Automation, Power and Signal Processing (CAPS), pp. 1–6
24. Bhatti UA, Yu Z, Yuan L, Zeeshan Z, Nawaz SA, Bhatti M et al (2020) Geometric algebra applications ingeospatial artificial intelligence and remote sensing image processing. IEEE Access 8:155783–155796
25. Bolourian Haghighi B, Taherinia AH, Mohajerzadeh AH (2018) "TRLG: fragile blind quad watermarking for image tamper detection and recovery by providing compact digests with quality optimized using LWT and GA," arXiv e-prints, arXiv-1803
26. "Bossbase dataset" (n.d.) https://www.kaggle.com/lijiyu/bossbase
27. Byrnes O, La W, Wang H, Ma C, Xue M, Wu Q (2021) " Data hiding with deep learning: A survey unifying digital watermarking and steganography," arXiv preprint arXiv:2107.09287
28. Camacho C, Kai W (2021) A comprehensive review of deep-learning-based methods for image forensics. J Imaging 7(4):69
29. Cao F, An B, Wang J, Ye D, Wang H (2017) Hierarchical recovery for tampered images based on watermark self-embedding. Displays 46:52–60

30. Cao H, Hu F, Sun Y, Chen S, Su Q (2022) Robust and reversible color image watermarking based on DFT in the spatial domain. Optik 169319:262
31. "Cassia-v2.0 Dataset:," (n.d.) https://www.kaggle.com/divg07/casia-20-image-tampering-detection-dataset
32. Celik MU, Sharma G, Saber E, Tekalp AM (2002) Hierarchical watermarking for secure image authentication with localization. IEEE Trans Image Process 11(6):585–595
33. Castro M, Ballesteros, DM, Renza D (2020) A dataset of 1050-tampered color and grayscale images (CG-1050). Data in Brief 28:104864. https://www.kaggle.com/saurabhshahane/cg1050
34. Chalamala SR, Kakkirala K. R (2015) "Local binary patterns for digital image watermarking," 2015 3rd international conference on artificial intelligence, modelling and simulation (AIMS), pp. 159-162
35. Chang YJ, Wang RZ, Lin JC (2009) A sharing-based fragile watermarking method for authentication and self-recovery of image tampering. EURASIP Journal on Advances in Signal Processing 2008:1–17
36. Chang CC, Lu TC, Zhu ZH, Tian H (2018) An effective authentication scheme using DCT for Mobile devices. Symmetry 10(1):13
37. Charkari NM, Chahooki MAZ (2007) " A robust high capacity watermarking based on DCT and spread spectrum," In 2007 IEEE International Symposium on Signal Processing and Information Technology. IEEE., pp. 194–197
38. Chaughule SS, Megherbi DB (2019) "A robust, non-blind high capacity & secure digital watermarking scheme for image secret information, authentication and tampering localization and recovery via the discrete wavelet transform," 2019 IEEE international symposium on Technologies for Homeland Security (HST).IEEE, pp. 1-5
39. Chen J, Kang X, Liu Y, Wang ZJ (2015) Median filtering forensics based on convolutional neural networks. IEEE Signal Process Lett 22(11):1849–1853
40. Coronel SLG, Ramírez BE, Mosqueda MAA (2016) Robust watermark technique using masking and Hermite transform. SpringerPlus 5(1):1–20
41. Cozzolino D, Poggi G, Verdoliva L (2015) Efficient Dense-Field Copy–Move Forgery Detection. IEEE Trans Inf Forensic Secur 10(11):2284–2297
42. Eugene B (2021) Data breaches: most significant breaches of the year 2021. https://www.identityforce.com/blog/2021-data-breaches
43. Dobre RA, Preda RO, Marcu AE (2018) "Improved active method for image forgery detection and localization on Mobile devices," 2018 IEEE 24th international symposium for design and Technology in Electronic Packaging(SIITME). IEEE, pp. 255–260
44. Dogan S, Tuncer T, Avci E, Gulten A (2011) A robust color image watermarking with singular value decomposition method. Adv Eng Softw 42(6):336–346
45. Dong J, Wang W, Tan T (2013) Casia image tampering detection evaluation database. In: 2013 IEEE China summit and international conference on signal and information processing. IEEE, China, pp 422–426
46. Cao Q, Xu L (2019) Unsupervised greenhouse tomato plant segmentation based on self-adaptive iterative latent dirichlet allocation from surveillance camera. Agronomy 9(2): 91. https://www.researchgate.net/publication/331168576/figure/fig1/AS:727678764199949@1550503545624/Sub-bands-separated-by-a-three-level-dyadic-discrete-wavelet-transform-DWT.png
47. Elshoura SM, Megherbi DB (2013) Analysis of noise sensitivity of Tchebichef and Zernike moments with application to image watermarking. J Vis Commun Image Represent 24(5):567–578
48. Fang H, Zhang W, Zhou H, Cui H, Yu N (2018) Screen-shooting resilient watermarking. IEEE Trans Inf Forensics Secur 14(6):1403–1418
49. Prabukumar M, Agilandeeswari L, Ganesan K (2019) An intelligent lung cancer diagnosis system using cuckoo search optimization and support vector machine classifier. J Ambient Intell Human Comput 10:267–293
50. Fita A, Endebu B (2019) Watermarking colored digital image using singular value decomposition for data protection. J Math Stat Anal 127:964–9726
51. Gao H, Chen Q (2021) A robust and secure image watermarking scheme using SURF and improved artificial bee colony algorithm in DWT domain. Optik 242:166954
52. Gómez-Moreno H, Gil-Jiménez P, Lafuente-Arroyo S, López-Sastre R, Maldonado-Bascón S (2014) A salt and pepper noise reduction scheme for digital images based on support vector machines classification and regression. Sci World J 2014:826405
53. Sunny S, Agilandeeswari L (2013) Secure data sharing of patient record in cloud environment using attribute based encryption. Int J Appl Eng Res 8(19)
54. "Growth Rate of Facebook," (n.d.) https://cdn.statcdn.com/Infographic/images/normal/10047.jpeg

55. Guo JM, Prasetyo H (2014) Security analyses of the watermarking scheme based on redundant discrete wavelet transform and singular value decomposition. AEU-Int J Electron Commun 68(9):816–834
56. Guo JM, Prasetyo H (2014) False-positive-free SVD-based image watermarking. J Vis Commun Image Represent 25(5):1149–1163
57. Hamidi M, El Haziti M, Cherifi H, El Hassouni M (2018) Hybrid blind robust image watermarking technique based on DFT-DCT and Arnold transform. Multimed Tools Appl 77(20):27181–27214
58. Han B, Du J, Jia Y, Zhu H (2021) Zero-watermarking algorithm for medical image based on VGG19 deep convolution neural network. J Health Eng 2021
59. Hartung F, Girod B (1998) Watermarking of uncompressed and compressed video. Signal Process 66(3):283–301
60. Hatami E, Rashidy Kanan H, Layeghi K, Harounabadi A (2023) An optimized robust and invisible digital image watermarking scheme in Contourlet domain for protecting rightful ownership. Multimed Tools Appl 82(2):2021–2051
61. Hongbo BI, Xueming LI, Zhang Y (2013) A novel HVS-based watermarking scheme in contourlet transform domain. Telkomnika Indonesian J Electr Eng 11(12):7516–7524
62. Hoshi AR, Zainal N, Ismail M, Rahem AART, Wadi SM (2021) A robust watermark algorithm for copyright protection by using 5-level DWT and two logos. Indonesian J Electric Eng Comput Sci 22(2):842–856
63. Huang Y, Lu W, Sun W, Long D (2011) Improved DCT-based detection of copy-move forgery in images. Forensic Sci Int 206(1–3):178–184
64. Hurrah NN, Parah SA, Loan NA, Sheikh JA, Elhoseny M, Muhammad K (2019) Dual watermarking framework for privacy protection and content authentication of multimedia. Futur Gener Comput Syst 94:654–673
65. Islam SM, Debnath R, Hossain S. A (2007) "DWT based digital watermarking technique and its robustness on image rotation, scaling, JPEG compression, cropping, and multiple watermarking," 2007 international conference on information and communication technology. IEEE., pp. 246-249
66. Issa M (2018) "Digital image watermarking performance improvement using bio-inspired algorithms," In: Hassanien, A., Oliva, D. (eds) Advances in Soft Computing and Machine Learning in Image Processing, Advances in Soft Computing and Machine Learning in Image Processing Studies in Computational Intelligence,730. Springer, Cham., vol. 730
67. Jana M, Jana B, Joardar S (2022) Local feature based self-embedding fragile watermarking scheme for tampered detection and recovery utilizing AMBTC with fuzzy logic. J King Saud Univ-Comput Inform Sci 34(10):9822–9835
68. Jayamalar T, Radha V (2010) Survey on digital video watermarking techniques and attacks on watermarks. Int J Eng Sci Technol 2(12):6963–6967
69. Jeffry B, Mammi H (2017) "A study on image security in social media using digital watermarking with metadata.," In 2017 IEEE conference on application, Information and Network Security (AINS) IEEE, pp 118–123
70. Jyothika A, Geetharanjin PR (2018) "Robust watermarking scheme and tamper detection using integer wavelet transform," 2018 2nd international conference on trends in electronics and informatics (ICOEI), pp. 676-679, May
71. Kang H, Leng L, Kim BG (2022) Data hiding of multicompressed images based on Shamir threshold sharing. Appl Sci 12(19):9629
72. Kessler B (2002) Constructions of orthogonal and biorthogonal scaling functions and multiwavelets using fractal. Adv Imaging Electron Phys 124:195–252
73. Kim C, Yang CN (2021) Self-embedding fragile watermarking scheme against tampering image by using AMBTC and OPAP approaches. Appl Sci 11(3):1146
74. Kim C, Shin D, Yang C, Leng L (2021) Data hiding method for color AMBTC compressed images using color difference. Appl Sci 11(8):3418
75. Kim C, Yang CN, Baek J, Leng L (2021) Survey on data hiding based on block truncation coding. Appl Sci 11(19):9209
76. "Kinetics datasets:" (n.d.) https://paperswithcode.com/dataset/kinetics-700
77. Korus P, Huang J Multi-Scale Analysis Strategies in PRNU-Based Tampering Localization. IEEE Trans Inf Forensic Secur 12(4):809–824
78. Kourkchi H, Ghaemmaghami S (2008) Image adaptive semi-fragile watermarking scheme based on RDWT-SVD. In: 2008 International Conference on Innovations in Information Technology. IEEE, pp 130–134

79. Lancini R, Mapelli F, Tubaro S (2002) " A robust video watermarking technique in the spatial domain," In International symposium on VIPromCom video/image processing and multimedia communications. IEEE
80. Laouamer L (2022) New informed non-blind medical image watermarking based on local binary pattern
81. Laouamer L, AlShaikh M, Nana L, Pascu AC (2015) Robust watermarking scheme and tamper detection based on threshold versus intensity. J Innov Digit Ecosyst 2(1–2):1–12
82. Lee GJ, Yoon EJ, Yoo KY (2008) A new LSB based digital watermarking scheme with random mapping function. In: 2008 International Symposium on Ubiquitous Multimedia Computing. IEEE, pp 130–134
83. Lefèvre P, Carré P, Fontaine C, Gaborit P, Huang J (2022) Efficient image tampering localization using semi-fragile watermarking and error control codes. Signal Process 190:108342
84. Li W, Yu N (2010) "Rotation robust detection of copy-move forgery," 2010 IEEE International Conference on Image Processing, pp. 2113–2116
85. Lin CH, Liu JC, Shih CH, Lee YW (2008) " A robust watermark scheme for copyright protection," 2008 International Conference on Multimedia and Ubiquitous Engineering (mue 2008) IEEE, pp. 132–137
86. Luo H, Yu FX, Huang ZL, Lu ZM (2011) Blind image watermarking based on discrete fractional random transform and subsampling. Optik 1:311–316
87. Maheshwari JP, Kumar M, Mathur G, Yadav RP, Kakerda RK.(2015) Robust digital image watermarking using DCT based pyramid transform via image compression. In: 2015 International conference on communications and signal processing (ICCSP). IEEE, pp 1059–1063
88. Maji P, Pal M, Ray R, Shil R (2020) "Image tampering issues in social media with proper detection," 2020 8th international conference on reliability, IEEE Infocom technologies and optimization (trends and future directions)(ICRITO), pp. 1272-1275
89. Manjunatha S, Patil MM (2021) Deep learning-based Technique for Image Tamper Detection. In: 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV). IEEE, pp 1278-1285
90. Mo J, Ma ZF, Huang QL (2012) An adaptive watermarking scheme using SVD in Contourlet domain. Adv Inf Sci Serv Sci 4(15):221–232
91. Mobasseri BG (2000) "A spatial digital video watermark that survives MPEG," Proceedings International Conference on Information Technology: Coding and Computing (Cat. No.PR00540) , pp. 68–73
92. Mohanrajan SN, Loganathan A (2022) Novel vision transformer–based bi-LSTM model for LU/LC prediction—Javadi Hills. Appl Sci 12(13):6387
93. Molina J, Ponomaryov V, Reyes R, Cruz C (2019) "Watermarking-based self-recovery and authentication framework for colour images," 2019 7th international workshop on biometrics and forensics (IWBF), pp. 1-6
94. Moltisanti M, Paratore A, Battiato S, Saravo L (2015) Image manipulation on facebook for forensics evidence. In: Image Analysis and Processing—ICIAP 2015: 18th International Conference, Genoa, Italy, September 7-11, 2015, Proceedings, Part II 18. Springer International Publishing, Italy, pp 506–517
95. Mun, S. M., Nam, S. H., Jang, H., Kim, D., & Lee, H. K. (2019). Finding robust domain from attacks: A learning framework for blind watermarking. Neurocomputing, 337, 191-202.
96. Munir R, Harlili H (2020) Application of Chaos-Based Fragile Watermarking to Authenticate Digital Video. In: Digital Forensic Science. IntechOpen
97. Ng T, Chang S, Sun Q (2004) Colombia gray: a data set of authentic and spliced image blocks. Columbia University, ADVENT Technical Report, 4
98. Ng TT, Hsu J, Chang SF (2009) Columbia image splicing detection evaluation dataset. DVMM lab. Columbia Univ Cal Photos Digit Libr. "Colombia color dataset:," (n.d.). https://www.ee.columbia.edu/ln/dvmm/downloads/AuthSplicedDataSet/AuthSplicedDataSet.htm
99. NR NR, Shreelekshmi R (2022) Fragile watermarking scheme for tamper localization in images using logistic map and singular value decomposition. J Visual Commun Image Represent 85:103500
100. D. T. Nguyen, Z. Zong, P. Ogunbona and W. Li, "Object detection using Non-Redundant Local Binary Patterns," 2010 IEEE International Conference on Image Processing, Hong Kong, China, 2010, pp. 4609-4612, doi: 10.1109/ICIP.2010.5651633."Local binary pattern image," (n.d.) no. https://ckyrkou.medium.com/object-detection-using-local-binary-patterns-50b165658368

101. Patel, M., Sajja, P. S., & Sheth, R. K. (2013). Analysis and survey of digital watermarking techniques. International Journal of Advanced Research in Computer Science and Software Engineering, 3(10), 1-15.
102. Patil RD, Metkar S (2015) "Fragile video watermarking for tampering detection and localization," in 2015 international conference on advances in computing, communications and informatics (ICACCI). IEEE., pp. 1661-1666
103. Plata M, Syga P (2020) " Robust spatial-spread deep neural image watermarking," In 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). IEEE., pp. 62–70
104. Prakash C, Kumar A, Maheshkar SEA (2018) An integrated method of copy-move and splicing for image forgery detection. Multimed Tools Appl 77:26939–26963
105. Preda RO (2014) Self-recovery of unauthentic images using a new digital watermarking approach in the wavelet domain. In: 2014 10th international conference on communications (COMM). IEEE, pp 1–4
106. Rafigh M, Moghaddam ME (2010) A robust evolutionary based digital image watermarking technique in DCT domain. In: 2010 Seventh International Conference on Computer Graphics, Imaging and Visualization. IEEE, pp 105–109
107. Rakhmawati L (2018) "Image fragile watermarking with two authentication components for tamper detection and recovery," in 2018 international conference on intelligent autonomous systems (ICoIAS). IEEE, pp. 35–38
108. Rao Y, Ni J (2016) A deep learning approach to detection of splicing and copy-move forgeries in images. In: 2016 IEEE international workshop on information forensics and security (WIFS). IEEE, pp 1–6
109. Rezaei M, Taheri H (2022) Digital image self-recovery using CNN networks. Optik 264:169345
110. Rhayma AH, Makhloufi, HH, Hmida AB (2018) "Semi fragile watermarking scheme for image recovery in wavelet domain," 2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), pp. 1–5
111. Sawant S.S, Manoharan P, Loganathan A (2021) Band selection strategies for hyperspectral image classification based on machine learning and artificial intelligent techniques–Survey. Arab J Geosci 14:1–10
112. Saini P, Ahuja R, Kaur A (2021) A review on video authentication technique exploiting watermarking methods. In: 2021 9th international conference on reliability Infocom technologies and optimization (trends and future directions)(ICRIT). ICRIT, pp 1–5
113. Sang J, Liu Q, Song CL (2020) Robust video watermarking using a hybrid DCT-DWT approach. J Electron Sci Technol 18(2):100052
114. Sawant SS, Prabukumar M, Loganathan A, Alenizi FA, Ingaleshwar S (2022) Multi-objective multiverse optimizer based unsupervised band selection for hyperspectral image classification. Int J Remote Sens 43(11):3990–4024
115. Scheibenreif L, Hanna J, Mommert M, Borth D (2022) "Self-supervised vision transformers for land-cover segmentation and classification," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1422–1431
116. Setiadi DRIM (2020) PSNR vs SSIM: imperceptibility quality assessment for image steganography. Multimed Tools Appl 80(6):8423–8444
117. Sharma V, Gangarde M, Oza S (2019) A spatial domain based secure and robust video watermarking technique using modified LSB and secret image sharing. ICTACT J Image Vid Process 10(1):2061–2070
118. Shukla D, Sharma M (2018) A novel scene-based video watermarking scheme for copyright protection. J Intell Syst 27(1):47–66
119. Singh B, Sharma MK (2021) Tamper detection technique for document images using zero watermarking in wavelet domain. Comput Electric Eng 89:106925
120. Singh B, Sharma DK (2021) SiteForge: Detecting and localizing forged images on microblogging platforms using deep convolutional neural network. Comput Industri Eng 162:107733
121. Sinhal R, Ansari IA, Ahn CW (2020) Blind image watermarking for localization and restoration of color images. IEEE Access 8:200157–200169
122. SIPI dataset: Allan Weber, 213-740-4147 (n.d.). https://sipi.usc.edu/database/database.php?volume=misc.
123. "Social Media Growth," (n.d.) https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/
124. Song C, Sudirman S, Merabti M, Al-Jumeily D (2011) "Region-Adaptive Watermarking System and Its Application," 2011 Developments in E-systems engineering. IEEE, pp 215–220

125. Soni B, Das PK, Thounaojam DM (2017) CMFD: a detailed review of block based and key feature based techniques in image copy-move forgery detection. IET Image Process 12(2):167–178
126. Soppari K, Chandra NS (2020) Development of improved whale optimization-based FCM clustering for image watermarking. Comput Sci Rev 37:100287
127. Standard test dataset-SIPI (n.d.). https://www.imageprocessingplace.com/root_files_V3/image_databases.htm
128. Sun W, Zhou J, Li Y, Cheung M, She J (2020) Robust high-capacity watermarking over online social network shared images. IEEE Trans Circ Syst Vid Technol 31(3):1208–1221
129. Tan L, He Y, Wu F, Zhang D (2020) A blind watermarking algorithm for digital image based on DWT. J Phys: Confer Ser 1518(1):012068
130. Tang W, Tan SLB, Huang J (2017) Automatic steganographic distortion learning using a generative adversarial network. IEEE Signal Process Lett 24(10):1547–1551
131. Thakur R, Rohilla R (2020) Recent advances in digital image manipulation detection techniques: A brief review. Forens Sci Int 312:110311
132. "The Copy-Move Forgery Database with Similar but Genuine Objects (COVERAGE) accompanies the following publication: COVERAGE– A NOVEL DATABASE FOR COPY-MOVE FORGERY DETECTION," (2016) IEEE International Conference on Image processing (ICIP)
133. Tohidi F, Paul M, Hooshmandasl MR (2021) Detection and recovery of higher tampered images using novel feature and compression strategy. IEEE Access 9:57510–57528
134. Tralic D, Zupancic I, Grgic S, Grgic M (2013) CoMoFoD—New database for copy-move forgery detection. In: Proceedings ELMAR-2013. IEEE, "CoMoFoD Dataset:," (n.d.), pp 49–54. https://www.vcl.fer.hr/comofod/examples.html
135. Tsai MJ, Chien CC (2008) "A wavelet-based semi-fragile watermarking with recovery mechanism," in 2008 IEEE international symposium on circuits and systems (ISCAS). IEEE, pp. 3033-3036
136. Vahedi E, Lucas C, Zoroofi RA, Shiva M (2007) "A new approach for image watermarking by using particle swarm optimization," 2007 IEEE International Conference on Signal Processing and Communications, pp. 1383–1386
137. Vassaux PB, Nguyen S, Baudry PB, Chassery J (2002) "Scrambling technique for video object watermarking resisting to MPEG-4," International Symposium on VIPromCom Video/Image Processing and Multimedia Communication, pp. 239–244
138. Venu KN, Sujatha BK (2021) Enhanced block based copy paste image forgery detection. Mater Today:Proc 2021. https://doi.org/10.1016/j.matpr.2021.01.189
139. Verma VS, Jha RK, Ojha A (2015) Digital watermark extraction using support vector machine with principal component analysis based feature reduction. J Vis Commun Image Represent 31:75–85
140. Verma V, Srivastava VK, Thakkar F (2016) "DWT-SVD based digital image watermarking using swarm intelligence," 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 3198–3203
141. "Video Dataset:," (n.d.) https://paperswithcode.com/datasets?mod=videos
142. "VIPP dataset:," (n.d.) http://clem.dii.unisi.it/~vipp/datasets.html
143. Wan W, Wang J, Zhang Y, Li J, Yu H, Sun J (2022) A comprehensive survey on robust image watermarking. Neurocomputing 448:226–247
144. Wang X, Wang J, Peng H (2009) "A semi-fragile image watermarking resisting to JPEG compression," in 2009 international conference on management of e-commerce and e-government. IEEE., pp. 498-502
145. Wang XY, Jiao LX, Wang XB, Yang HY, Niu PP (2018) A new keypoint-based copy-move forgery detection for color image. Appl Intell 48(10):3630–3652
146. Wang XY, Liu YN, Xu H, Wang P, Yang HY (2018) Robust copy–move forgery detection using quaternion exponent moments. Pattern Anal Applic 21(2):451–467
147. Xu H, Kang X, Chen Y, Wang Y (2019) Rotation and scale invariant image watermarking based on polar harmonic transforms. Optik 183:401–414
148. Yao B, Jiang X, Khosla A, Lin AL, Guibas L, Fei-Fei L (2011) Human action recognition by learning bases of action attributes and parts. In: 2011 International conference on computer vision. IEEE, pp 1331–1338
149. Yeo IK, Kim HJ (2003) Generalized patchwork algorithm for image watermarking. Multimed Syst 9(3):261–265
150. Yu C (2020) "Attention based data hiding with generative adversarial networks," in Proceedings of the AAAI conference on artificial intelligence, Vol. 34, No. 01, pp. 1120–1128
151. Yu X, Wang C, Zhou X (2017) Review on semi-fragile watermarking algorithms for content authentication of digital images. Future Int 9(4):56

152. Zampoglou M, Papadopoulos S, Kompatsiaris Y (2015) " Detecting image splicing in the wild (WEB)," IEEE International Conference on Multimedia & Expo Workshops (ICMEW)
153. Zhang Y, Thing VL (2017) A multi-scale noise-resistant feature adaptation approach for image tampering localization over Facebook. In: 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP). IEEE, pp 272–276
154. Zhang X, Cui L, Shao L (2012) "A fast semi-fragile watermarking scheme based on quantizing the weighted mean of integer Haar wavelet coefficients," in 2012 symposium on photonics and optoelectronics. IEEE, pp. 1-4
155. Zhang H, Wang C, Zhou X (2017) Fragile watermarking for image authentication using the characteristic of SVD. Algorithms 10(1):27
156. Zheng PP, Feng J, Li Z, Zhou MQ (2014) A novel SVD and LS-SVM combination algorithm for blind watermarking. Neurocomputing 142:520–528
157. Zhou G, Lv D (2011) An overview of digital watermarking in image forensics. In: 2011 fourth international joint conference on computational sciences and optimization. IEEE, Kunming and Lijiang City, China, pp 332–335
158. Zhou X, Ma J, Du W (2013) "SoW: a hybrid DWT-SVD based secured image watermarking," In PROCEEDINGS OF 2013 International Conference on Sensor Network Security Technology and Privacy Communication System ,IEEE, pp. 197–200
159. Zhou N, Luo A, Zou W (2019) Secure and robust watermark scheme based on multiple transforms and particle swarm optimization algorithm. Multimed Tools Appl 78:2507–2523
160. Zhu J, Kaplan R, Johnson J, Fei-Fei L (2018) "Hidden: Hiding data with deep networks," In Proceedings of the European conference on computer vision (ECCV), pp. 657–672
161. Zigomitros A, Papageorgiou A, Patsakis C (2012) Social network content management through watermarking. In: 2012 IEEE 11th international conference on trust, security and privacy in computing and communications. IEEE, Liverpool, UK, pp 1381–1386
162. Zong T, Xiang Y, Natgunanathan I, Guo S, Zhou W, Beliakov G (2015) Robust histogram shape-based method for image watermarking. Circuits and Systems for Video Technology. IEEE Trans 25:717–729