Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Exploring Watermarking Strategies in Deep Learning: A Comprehensive Review

Hadeel Mohsen Ibrahim[1], Methaq Talib Gaata[2], Huda Abdulaali Abdulbaqi[3]

1 Department of Computer Science, College of Science, Al-Mustansiriyah University, Al-Waziriyah, Baghdad, Iraq, hadil.muhsen@uomustansiriyah.edu.iq

2 Department of Computer Science, College of Science, Al-Mustansiriyah University, Al-Waziriyah, Baghdad, Iraq, dr.methaq@uomustansiriyah.edu.iq

3 Department of Computer Science, College of Science, Al-Mustansiriyah University, Al-Waziriyah, Baghdad, Iraq. huda.it@uomustansiriyah.edu.iq

ARTICLE INFO

Article history:
Received: 15 /03/2025
Rrevised form: 22 /04/2025
Accepted : 30 /04/2025
Available online: 30/06/2025

Keywords:

image watermarking , deep learning, Generative adversarial network (GAN).

ABSTRACT

Recently, there has been a notable increase in the need for the generation, dissemination, and storage of extensive volumes of multimedia data, especially digital images, generated by different intelligent devices and sensors. Along with other security vulnerabilities, such activity results in unauthorized access and false use of data. Embedding a watermark design into a digital cover and then removing it helps to address ownership conflicts and copyright infringement concerns related to the media data. Deep-learning methods are currently rather helpful for watermarking because of their high accuracy, great precision, and strong learning ability. This article presents a comprehensive examination of watermarking techniques used in deep learning contexts. We start by explaining the basic ideas of both traditional and learning-based digital watermarking, and then we look at common watermarking methods that use deep learning models. We then provide a succinct summary and comparison of the most recent contributions in the literature. Finally, we highlight the challenges related to obfuscation and suggest avenues for further study.

MSC..

https://doi.org/10.29304/jqcsm.2025.17.22187

Introduction

Recent years have seen a notable increase in the need for the generation, dissemination, and preservation of large volumes of multimedia data, namely digital images, derived from diverse intelligent devices and sensors. This poses difficulties, such as unauthorized access and deceitful use of this data, in addition to other security concerns. Watermarking is the systematic integration of a watermark design into a digital cover, which is then extracted to resolve ownership conflicts and issues about copyright infringement concerning the media data. At present, deep-learning techniques provide significant benefits in watermarking due to their exceptional accuracy, unparalleled precision, and resilient learning capabilities. In this paper, we provide a comprehensive examination of watermarking techniques applicable in deep-learning environments. Commencing with the core ideas of conventional and learning-based digital watermarking, we next analyze the extensively used deep-learning model-based digital watermarking methodologies. The next part briefly summarizes and compares the most recent contributions in the literature. At last, we discuss the difficulties with obfuscation and propose some possible directions of research figure 1 show contemporary implementations of watermarking.

The primary objective of watermarking methods is to enhance three essential criteria: The primary objective of watermarking methods is to enhance three essential criteria figure2:

Imperceptibility: Imperceptibility is the ability of a watermarking technique to include data in digital media without clearly changing the original content. To maintain the integrity and use of the medium, the watermark must remain hidden or indistinguishable to humans. This criterion is crucial, as a clearly identifiable watermark may be changed

or deleted, therefore compromising the security of the protected content. The effectiveness of a watermarking technique is determined most of all by its capacity to maintain the visual or auditory integrity of the host media. Advanced learning-based watermarking systems use deep neural networks to identify suitable embedding locations, thereby ensuring subtle and undetected changes. Thus, imperceptibility ensures that the watermark preserves the integrity of its visual or aural features without interfering with the intended experience of the medium.

Capacity: Capacity in watermarking is the volume of information that may be integrated into a digital medium, such as an image, video, or audio file, without compromising its quality. It denotes the volume of data that the watermarking technique can transmit while maintaining imperceptibility and resilience. An increased capacity facilitates the incorporation of more intricate watermarks, such as copyright details, identifying codes, or security elements. However, increasing capacity can make it harder to keep the watermark hidden and strong, because adding more data might lower the quality of the media or make the watermark easier to see and remove. Optimizing capacity entails achieving a balance between the quantity of embedded data and the finesse of its integration into the medium. Learning-based watermarking systems use deep learning models to find the best balance, automatically changing how data is added to improve capacity while making changes less noticeable.

Robustness: Robustness means the watermark's ability to survive different types of attacks, whether they are intentional or accidental, like compression, adding noise, cropping, scaling, and filtering, while still being recognizable. A robust watermark in watermarking is one that can be reliably retrieved or confirmed, even after the host medium has been altered or subjected to harmful interference. Robustness is essential, since the primary objective of watermarking is to safeguard digital material and establish ownership. If the watermark is readily removed or modified, the security and integrity of the watermarking system are undermined. Deep learning-based watermarking methods improve resilience by acquiring resilient embedding patterns via training on varied datasets. These algorithms adjust to the content attributes, ensuring that the watermark is implanted in areas of the media that are less prone to deterioration and therefore preserve its integrity under diverse assault scenarios.

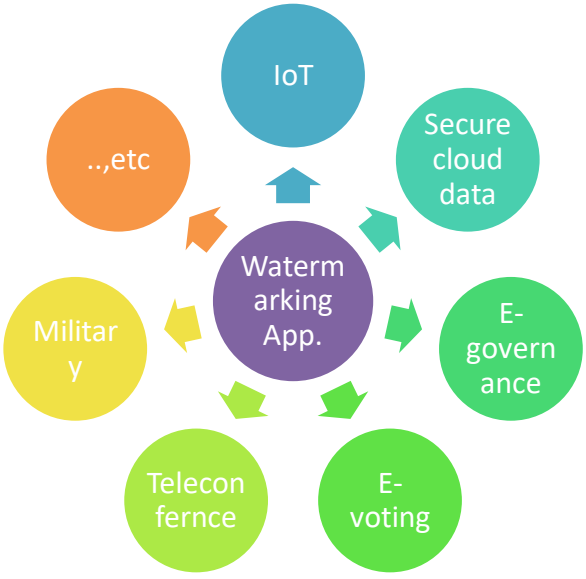

Fig. 1 Contemporary implementations of watermarking.

The conventional technique of watermarking involves embedding an authentication code or secret key into an image to verify its authenticity before transmission over a public network, such as the internet. This encoded data functions as a unique identifier or proof of ownership, ensuring that the image remains unaltered throughout transmission or storage. The methodology typically consists of two fundamental phases: embedding and extraction.

In the embedding step, a watermark—possibly a private key, logo, or verification code —is integrated into the original picture. This watermark is often embedded in certain areas or characteristics of the picture, ensuring it

stays undetectable to human viewers while maintaining the image's visual integrity. The watermarked picture is then transferred or shared.

During the extraction step, the receiver retrieves the embedded watermark via a designated methodology and validates it against the original secret key or authentication code. If the recovered watermark matches the expected key, we validate the picture as legitimate. This procedure aids in identifying any illicit modifications, so safeguarding the integrity and validity of the digital asset. Figure 3 would generally depict this two-step procedure, emphasizing the transition from embedding to verification.

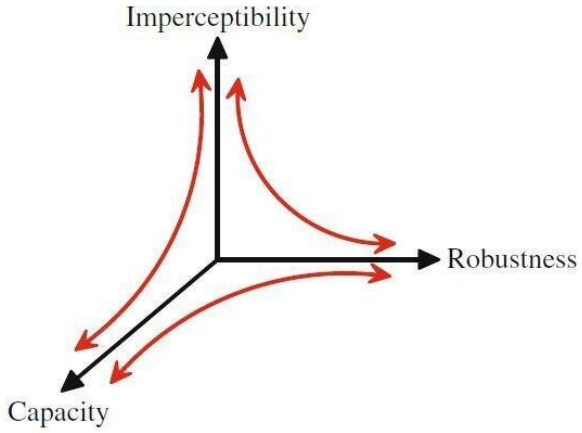

Fig. 2 Trade-off Triangle in Watermarking.

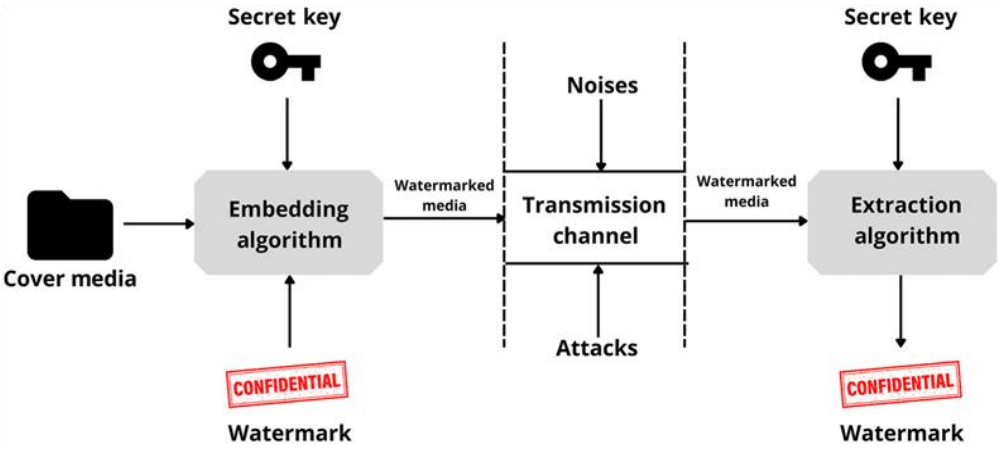

Fig. 3 Summary of watermarking architecture.

Deep learning-based watermarking methods have attracted a lot of attention lately because of their improved performance over conventional methods. These technologies are a great choice for watermarking since they provide several noteworthy advantages. Deep learning makes it easier to create watermarks that are more resilient and guarantee detectability even in the face of many attacks or changes. Second, deep learning models are excellent at determining the best embedding location inside the cover material, improving the watermark's positioning without lowering the caliber of the content. Third, by finding the right embedding strength, these models can make sure the watermark is hidden well and can withstand changes, balancing how strong it is with how invisible it appears.

Additionally, deep learning approaches may also mimic media assaults, thus aiding in the development of more reliable watermark extraction algorithms. Moreover, they contribute to accuracy by eliminating erroneous extractions and denoising watermarking components. All these advantages come with a complication in watermarking based on deep learning: models are susceptible to insecurity and privacy since the information that models contain must be protected from unauthorized access. Therefore, because useful information can be embedded within devices and models using deep learning watermarking methods, they have become crucial not only in settling ownership disputes but also in copyright infringement cases. Moreover, they constitute proof of ownership for authentication purposes.Numerous studies on the safeguarding of digital content, devices, and artificial constructs by watermarking techniques have been published recently. References 1, 10, 14, and 15 provide an extensive examination of many aspects of watermarking in digital security. Additionally, Ref. 1 gives a brief overview of the components and uses of deep-learning algorithms in different stages of watermarking techniques, showing how they can improve strength and flexibility. Reference 10 examines several watermarking approaches within the realm of artificial intelligence, evaluating their applications for safeguarding AI models. Reference 14 looks closely at how watermarking can protect deep-learning algorithms, especially in terms of proving who owns a model and preventing copyright issues. Particularly with an eye on the employment of deep-learning techniques to safeguard digital information and artificial intelligence assets, Reference 15 provides a thorough analysis of intellectual property protection.Our main goal is to deliver a complete study of the fundamental elements of deep-learning algorithms often used in watermarking for possession verification, right protection, and algorithm security. This work expands on the above papers by assessing, in the area of watermarking, the applicability and functionality of notable deep-learning models like CNNs, GANs, RNNs, and autoencoders. This paper presents a thorough review of many sophisticated methods for watermarking and data concealment; therefore, it provides important insights from several angles.

2. Learning-based Watermarking

Using deep-learning architectures to improve watermark embedding and extraction, learning-based watermarking has emerged as a successful approach for protecting digital data. Conventional watermarking methods are less versatile and robust in various media environments, as occasionally they need manually created feature representation. In contrast, deep-learning models can automatically create layered data representations from raw input like images or audio, so they don't need people to define features manually. Deep neural networks (DNNs), recurrent neural networks (RNNs), convolutional neural networks (CNNs), and generative adversarial networks (GANs) can all be used to effectively add watermarks to media content by using this ability. Thanks to their outstanding efficiency in image processing applications, convolutional neural networks (CNNs) have become very popular. Recent studies indicate that deep-learning algorithms provide great adaptability in selecting suitable embedding locations and strengths as well as enhance the imperceptibility and robustness of watermarks. Deep network usage in watermarking has various applications, including secure embedding, accurate extraction, and watermark generation. These models have demonstrated outstanding resilience against many attacks, including compression, noise, and geometric distortions, thereby enhancing digital rights management and copyright protection by imitating human cognitive learning capacities.

3. Exploring Deep Learning in Contemporary Watermarking Techniques

Recent studies have shown different aspects of deep learning, such as convolutional neural networks, generative adversarial networks, and deep neural networks; how they can be used for watermarking; and the methods for adding and removing watermarks while ensuring they are strong and not noticeable. While GANs generate undetectable but strong watermarks, CNNs are useful for determining ideal embedding sites. The watermarking method benefits from deep neural networks increasing complexity. This part will examine how the models operate in a sophisticated manner and assess studies providing some understanding of the contribution of the models and the development they propose to watermarking approaches.

3.1 CNN-based Watermarking

Convolutional Neural Networks (CNNs) are deep learning architectures characterized by their convolutional operations, inspired by the human brain's visual cortex. A convolutional neural network (CNN) typically has five layers: the input layer, convolutional layer using rectified linear unit (ReLU), pooling layer, fully connected layer, and output layer. The model accepts an image as input and uses convolutional filters to extract essential information. The ReLU activation function nullifies negative values by assigning them a value of zero and uses

pooling to decrease the quantity of limits. The summary data is sent to a fully linked layer for categorization. Convolutional neural networks (CNNs) are extensively used in picture watermarking and data concealment owing to their comparatively low complexity. The algorithms efficiently manage the processes of watermark embedding, extraction, visual quality improvement, and optimum embedding position detection. Dharwadkar and Ingaleshwar introduced an optimization-based watermarking method by means of deep convolutional neural networks (CNN) to embed the watermark picture into a cover image. They demonstrate the efficacy of CNNs in this field. Thus, convolutional neural networks (CNNs) are essential in contemporary watermarking methodologies. The idea entails a watermarking technique using a deep Convolutional Neural Network (CNN) for feature extraction and the incorporation of a watermark into a cover image. The process starts with the disaggregation of the cover image into many grids to identify significant features along the gridlines. This step facilitates the decomposition of the image into smaller, manageable portions, hence improving the accuracy of watermark placement. A deep convolutional neural network, optimized in several ways, is used to identify the most "interesting" regions in the image—suitable for watermark insertion while preserving image quality.In this method, identification of ideal locations for watermark insertion depends critically on fitness measures. Some wavelet sub-bands are chosen as they include less susceptible portions of the picture rich in information and less prone to deterioration during the embedding process. Recovering the watermark guarantees exact extraction of the pattern by use of ownership of the cover media and the previously calculated fitness value. In peak signal-to-noise ratio (PSNR), a measure of picture quality, the approach shows better performance than conventional methods mentioned in 19. Still, despite its strength, the method has not been fully investigated for its total embedding and recovery costs—including computing resources and time. More study is also required to evaluate its resistance against a wider spectrum of assaults at different noise levels. This emphasizes the necessity of a more thorough assessment of the efficiency of the technique in real-world scenarios where photos encounter different aberrations.

Using a deep convolutional neural network (CNN), the proposed watermarking method efficiently extracts features and inserts watermarks in a cover picture. The method begins by dividing the cover picture into small grids, which facilitates the identification of conspicuous features along the gridlines. The division helps determine suitable embedding sites and splits the picture into reasonable portions for further processing. By inserting watermarks in areas less noticeable to the human eye, the grid analysis helps the system maintain the picture quality. The "interesting" areas within the processed grids are then found using a deep convolutional neural network trained using several optimization methods. These are the perfect places for watermark insertion with the least effect on picture visual integrity. Using CNN, this method allows the model to learn and find the best spots for embedding by recognizing complex patterns and features that are usually strong against attacks or changes. …This method relies on fitness measurements to ascertain the ideal embedding placements. To make it stronger, specific wavelet sub-bands are selected because they contain important parts of the image that can handle different attacks (like compression or added noise). Embedding the watermark inside these sub-bands optimizes both resilience and imperceptibility. In the watermark recovery phase, the cover media and the pre-calculated fitness function are essential for the precise extraction of the embedded watermark, hence preserving the integrity of the watermarking process. The method shows a significant increase in peak signal-to-noise ratio (PSNR)—a key measure of image quality—compared to traditional watermarking techniques (reference 19). An elevated PSNR indicates that the encoded watermark has little influence on the image's perceived quality. Nonetheless, despite its encouraging efficacy, the approach has not been thoroughly examined regarding embedding and recovery expenses, including computing complexity and temporal demands. Additional study is required to examine its resilience against a broader range of assaults at varying noise levels. An extensive assessment of this technique's efficacy in practical applications is crucial, particularly in contexts where pictures encounter typical aberrations and intentional alterations.A blind watermarking solution for copyright protection, developed by Mun et al., aims to improve how strong and hidden the watermark is using deep convolutional neural networks. This method entails first partitioning both the carrier (cover image) and the watermark substance into non-overlapping segments. An encryption key is calculated using the watermark data associated with these block locations. According to this data, the watermarks are included in the transporter picture by means of a deep convolutional neural network machine model. Furthermore, once integrated, the model also enables the retrieval of the watermark from the cover picture, therefore guaranteeing a blind watermarking method (i.e., the extraction does not need the original image). This approach exhibits resilience against many types of assaults, including salt and pepper noise, while maintaining a high level of imperceptibility. However, a significant drawback is the greatly increased computational expense, which affects its efficiency.

Reference 35 introduced a watermarking approach based on encryption-compression. Using the lifting wavelet transform (LWT), randomized singular value decomposition (RSVD), and Heisenberg decomposition (HD), this approach employs transform domain watermarking. To further enhance the extracted watermark's quality, a CNN-

based denoising network is employed. This advanced approach outperformed present methods (reference 36), showing a 27.84% rise in resilience and accuracy when tested against many hazards. Using complex transformations in combination with CNN-based denoising offers a strong and safe watermarking technique, therefore highlighting the possibilities of deep learning in watermarking uses.Based on quaternion discrete cosine transform (QDCT), Hsu and Hu put forward a watermarking technique in the transform domain. This approach strikes a compromise between watermarked picture imperceptibility and robustness by using the gray wolf optimizer. We use a blind watermark extraction method in conjunction with a denoising CNN model to improve watermark quality. The results showed that this method is better than traditional methods (reference 41), indicating it is more resistant to attacks while keeping good visual quality. To maximize the watermarking process, Kandi et al. presented yet another transform domain watermarking system using a CNN network. This approach uses CNN's auto-encoder features to improve resilience by embedding the watermark via CNN weights, which resist several types of assaults. The embedding technique makes use of input-output information, therefore increasing the watermark's general lifetime.Nagai et al. offered a watermarking method including a secret into several convolution layer groups of the first network to safeguard CNN model ownership. While this strategy helps to reduce certain hazards, it is not effective in handling complex assaults like watermark rewriting and surrogate model attacks. Guan et al. suggested another watermarking technique aimed at securing CNN networks by using the principle of pruning in model compression. They embed the watermark data as a hash, implemented using the safe hash method SHA-256, into the convolution layer of a residual neural network (ResNet152). Although this approach ensures ownership protection, its high computational costs limit its applicability in real-time scenarios.

Table 1 Research Summary CNN-based Watermarking

Ref.	Method	Drawback	Result
[21]	DeepSigns: Generic watermarking framework for DL models	Limited to protecting the ownership of deep learning models	Provides a generic framework for embedding watermarks in deep learning models
[22]	Wavelet Transform Applications	Generalized focus on wavelets; lacks specific watermarking details	Introduces modern applications of wavelet theory
[23]	DWT-SVD and CNN-based robust image watermarking	The complexity of balancing durability and capacity	Achieves robustness and high capacity for image watermarking
[24]	CNN-based watermarking for smart city applications	Security-focused, but lacks detailed discussion on watermark embedding.	Provides a security-guaranteed image watermarking scenario for smart city applications
[25]	Compression-resistant model watermarking	Pressure resistance comes with arithmetic expenses	Introduces a watermarking technique resilient to model compression for IP protection
[26]	CNN-based watermarking adaptive to image resolution	Adaptive resolution limits compatibility across different scenarios	A digital image watermarking method adaptive to the resolution of both image and watermark

3.2 GAN - Watermarking

Generative adversarial networks (GANs) consist of two basic parts—the generator network and the discriminator network—and are a type of deep-learning model that is very good at creating new content. Using their unique generating and discriminative characteristics, GANs are quite useful in the field of watermark production and verification. From random noise, the generator network (G) creates a sample S designed to reproduce a target distribution P(X) following a specified distribution P(Z). This method challenges attacker prediction or removal of the watermark by allowing GANs to create original watermarks for every input, hence boosting watermarking

security. Conversely, we train the discriminator network (D) to differentiate fraudulently generated data from real data. Over the training process, the discriminator learns to detect the smallest differences between produced generator images and real watermarked ones. It penalizes itself for misclassifications, whether it falsely labels a real case as false or a false case as real. This adversarial training process improves the generator's ability to produce highly realistic and robust watermarks. As a result, GANs have become a powerful tool in watermarking, providing enhanced security and adaptability in both watermark generation and verification.Wu et al. developed a deep-learning algorithm intended to demonstrate photos distorted by dense watermarking, hence enhancing image recovery quality and verification performance. The model uses in a GAN-like architecture a generator and a discriminator. Originally transferring highly watermarked, damaged images to a representation vector, the generator operates as an autoencoder, then decodes this vector back into an RGB image. This procedure aids in the restoration of the picture's visual quality. Working to reduce feature loss, the discriminator controls the content quality of the produced pictures, therefore guaranteeing the authenticity of the recovered photos.The model utilizes ResNet-46 to extract features from both the ground truth images and the recovered images, enhancing the authentication process. At a false positive rate (FPR) of 1%, this method obtains an impressive verification accuracy of 96.29%, demonstrating its efficacy in differentiating between genuine and distorted photos. Another study (Ref. 50) proposed a robust data hiding method based on a GAN to safeguard original documents. Geometric fixes first help the text be in its intended shape. Thereafter, the adversarial network creates a watermarked document by embedding secret knowledge using a pseudo-random integer. This method offers better security for verifying document authenticity and shows more resilience than approaches mentioned in Ref. 51.

Wei et al. introduced a robust watermarking system based on variational autoencoder (VAE) networks for copyright protection. The architecture has three main subnetworks: an encoder, a decoder, and a detector, together known as the embedder and extractor networks. Throughout the training phase, an encoder inserts a low-bit watermark picture into the crowd image. Working together, the encoder and decoder create a powerful representational depiction of the cover picture, therefore guaranteeing secure watermark insertion and preserving picture quality. The system is specifically designed to accurately extract the low-bit watermark feature, as the watermarked picture serves as the detector subnetwork. This approach significantly improves the visual quality of the watermarked image, but further research is required to evaluate its resilience completely against different sorts of attacks.A separate study presented a blind watermarking approach using deep learning (Ref. 55). This approach is comprised of four primary elements: an encoder, a decoder, two identical noise layers, and an adversarial discriminator. The encoder and decoder embed and erase the watermark using the same noise levels to increase its resistance against many kinds of attacks. Moreover, the adversarial discriminator helps strengthen the resilience and obfuscation of the watermark, thereby guaranteeing its detection and removal resistance. Using deep learning to enhance the efficient embedding and watermark extraction provides a robust solution for protecting digital copyright.It is famous for its robustness against many attacks with excellent imperceptibility. However, its major drawback is high model complexity, which may impede its efficiency and practical application. Fan et al.20 found a DWI picture transmission in a 4.9-MB lossless format during the transmission, which occupied a little space in the text measurement but required storage space at both the viewing and transmitting ends. Fan et al.20 proposed a multiscale barely-aliasing watermarking method (Ref. 56) to mitigate security risks and discourage proprietary use cases. This method protects the DWI images by adding both multi-scale characteristics and a GAN. Using full-scale characteristics to reconstruct the DWI images, this method approaches the original images closely. This method embeds the watermarks without significantly altering the image's visual quality. We then merge them with the core multiscale rebuilt features, thereby reinforcing the elimination of any node attacks. We apply an enhanced boundary equilibrium GAN discriminator to improve the visual quality of the rebuilt image. This makes it unwanted, and the integrity of the watermarked photo remains intact. Additionally, the model utilizes exact properties of the watermark distribution, specifically multiscale max-pooling and pyramid filters. Exploiting this multiscale strategy enables the model to learn under diverse imagesfeatures, improving the invisibility and robustness of the watermark. … Fang et al. proposed a triple-phase watermarking strategy to reduce picture distortion, thus offering a robust technique for data hiding. This method consists of three stages: a noise-free stage, a mask-guided frequency augmentation stage, and an adversarial training stage. First, we train a full encoder-decoder network using a just-noticeable difference (JND) mask picture loss scheme. This process ensures that inserting the watermark does not greatly alter the image quality. The second stage uses frequency changes based on masks to adjust the encoded features, helping the watermark fit into different frequency areas of the image and making it stronger. In the third step, adversarial training is used to teach the decoder how to more effectively cope with non-differentiable distortions, thereby enhancing the watermark's resistance to attacks. Reference 58 has found this combined method to be more robust than others.

A further study (Ref. 62) introduced a semi-fragile watermarking framework for media authentication, grounded on deep learning methodologies. The approach has three elements: The approach consists of an aggressive discriminator algorithm, a decoder network, and an encoder network. The encoder embeds the watermark onto the input photographs. The watermarked picture then goes through two changes: a sample from a transformation set that is not harmful and another sample from a transformation set that is harmful, therefore producing both benign and malicious variations of the image. Thereafter, an adversarial network decodes the pictures so that it can distinguish benign ones from malicious ones. More research on the effectiveness and robustness of this approach for tamper detection will help us evaluate its sensitivity to a wider spectrum of image processing constraints.

Table 2 Research Summary GAN-based Watermarking

Ref.	Method	Drawback	Result
[27]	Supervised GAN watermarking for IP protection	That GAN models are computationally expensive	Introduces a supervised GAN framework for robust watermarking in intellectual property protection
[28]	Robust blind image watermarking based on interest points	Reliance on interest points affect performance in complex images	Achieves robust blind watermarking using interest points, resilient against various distortions
[29]	ARWGAN: Attention-guided GAN-based watermarking	Requires extensive training for optimal attention guidance	robust GAN model with attention mechanisms for enhanced image watermarking
[30]	Autoencoder-CNN based image watermarking	Autoencoders face challenges in highly complex image data	Utilizes autoencoder and CNN for efficient embedding and extraction of watermarks in images
[31]	Deep learning-based digital image watermarking	lack generalizability across diverse image types	Deep learning technique designed for improved image watermarking efficiency
[32]	DwiMark: Multiscale robust deep watermarking for DWI images	Focuses on specific medical images, limiting broader applicability	Provides a robust watermarking framework for diffusion-weighted imaging
[33]	Encoded feature enhancement in watermarking network	introduce additional complexity in real-scene distortions	Enhances encoded feature robustness to distortions encountered in real-world watermarking

3.3 *DNN - Watermarking*

Watermarking has excellent use in the area of Deep Neural Network (DNN) architecture. Between an input and an output layer, there are many hidden layers. These networks use a parametric function $p=F\delta(x)$ to convert raw inputs (like images or sounds) into an output, essentially simulating the working of the human brain. Network design and layered aggregated data help define the value of p. By means of parameter θ, the network seeks to reduce the backpropagation-based loss between anticipated and ground-truth labels. Deep neural networks (DNNs) may then encode and decode watermarks for uses like network security, embedding, and extraction after this training.The watermarking method recommended by Hou et al. is based on an improved version of multiple histogram modifications. Within this technique, the cover picture is divided into two sets of pixels, and the deep neural network (DNN) is used to produce many histograms using a classification methodology. Subsequently, the

watermark is included in the selected bins of these histograms. Even though our method works better than current methods (Ref. 64) in terms of peak signal-to-noise ratio (PSNR) and keeps great visual quality, we haven't thoroughly tested how well it stands up to various attacks and the overall cost of embedding.Zhang and colleagues (Ref. 67) introduced a watermarking methodology for copyright protection of deep neural network (DNN) models. This method uses a better threat model to allow for checking the system without seeing its inner workings and accessing the API, which solves the limitations of the Uchida et al. approach. Notwithstanding its commendable precision, the approach is deficient in a clear evaluation of its overhead and security. An alternative method for copyright protection of deep neural networks (DNN) was presented by Wu et al. (Ref. 68). Presented here is the watermarked picture generated from the output of the model, which can only be authenticated by a dedicated extraction network. Nevertheless, the assessment of this system is limited to three specific categories of assaults, and the reporting of its computing time lacks transparency, therefore limiting its use in real-time scenarios. The approach proposed by Deeba et al. (Ref. 69) involves the generation and integration of watermark data into the neural network to protect ownership. Verification entails the examination of certain input-output pairings, but its assessment is limited to two attack types and does not provide information on the execution time. Reference 70 introduced a deep neural network (DNN)-based watermarking method for verifying multimedia ownership. To integrate a 1-bit binary watermark into the DCT coefficients, the cover picture is partitioned into 8 × 8 pixel blocks. This method attains optimal Peak Signal-to-Noise Ratio (PSNR), resulting in almost indistinguishable

Table 3 Research Summary DNN-based Watermarking

Ref.	Method	Drawback	Result
[34]	Digital watermarking using deep neural network	face difficulties in highly variable image datasets	Utilizes deep neural networks for effective digital watermarking
[35]	Reversible data hiding based on multiple histograms modification and DNN	Reversibility reduce embedding capacity	Provides a reversible data hiding method, ensuring original data can be perfectly recovered
[36]	Intellectual property protection of DNNs with watermarking	vulnerable to specific watermark removal attacks	Introduces a method to protect intellectual property of deep neural networks via watermarking
[37]	DNN watermarking: Four challenges and a funeral	Challenges in scalability, robustness, and security	Discusses four critical challenges in DNN watermarking, highlighting the limitations in current approaches
[38]	Deep serial number: Computational watermarking for DNN IP protection	face computational overhead	computational watermarking technique for DNNs using deep serial numbers for IP protection
[39]	Digital watermarking using deep neural networks	face computational overhead	computational watermarking technique for DNNs using deep serial numbers for IP protection

3.4 Others DL

Besides the techniques mentioned earlier, watermarking methods have also used advanced learning models such as artificial neural networks (ANNs), backpropagation neural networks (BPNNs), and recurrent neural networks (RNNs). Sinhal et al. proposed a low-cost, blind watermarking method for digital color images. (Ref. 71) for the purpose of ownership verification and copyright protection. The picture undergoes conversion into the YCbCr model by using randomly selected 4 × 4 chunks of the Y-module. We create an embedded binary watermark by analyzing the selected component using integer wavelet transformation (IWT) and utilizing a cost-effective artificial neural network (ANN) model. Similarly, Islam et al. (Ref. 72) presented a dependable watermarking technique in the LWT (lifting wavelet transform) domain by using an artificial neural network (ANN). An artificial neural network

(ANN) model includes the watermark in the randomly generated coefficients of the LWT cover picture and then retrieves it. However reliable and blind, the method lacks a clear evaluation of embedding and extraction expenses and has restricted watermark capability for applied use.The suggested watermarking method utilizes edge detection by Kazemi et al. (Ref. 73) and involves the embedding of secret information inside the edges of a color picture. The edges of the RGB media are detected by an edge detector and then analyzed using contourlet transform to ascertain the maneuvering mechanisms. A genetic algorithm is used to mutate the logo picture to improve its verification before incorporating it into the cover image. Watermark extraction uses a hybrid approach that combines differential evolution with multilayer perceptron. Nevertheless, this method has little resistance to certain threats.Singh et al. (Ref. 74) developed a deep-learning-based watermarking method that embeds several watermarks into the DWT-DCT area of the cover picture. The method involves segmenting the cover photo using a three-level Discrete Wavelet Transform (DWT), focusing on the low-frequency (LL3) and low-high (LH2) bands to include the watermark. We implement selective encryption on the watermarked picture to conserve expenses, and a Bayesian Process Neural Network (BPNN) mitigates distortion effects on the retrieved watermark. While this approach offers data confidentiality and resilience against assaults on par with Ref. 73, a thorough examination of its security and cost is necessary. Singh et al. (Ref. 75) used a long short-term memory-based recurrent neural network (LSTM-RNN) to achieve data hiding. The method uses LSTM-RNN to minimize changes between the original and predicted signals and hides watermark data in the TP section of an ECG signal. Although this technique exceeds conventional methods (Ref. 72), its low watermark capacity limits its pragmatic use.Wang et al. (Ref. 76) presented a watermarking method grounded on blind DCT-SVD mapping. To boost the watermark's strength, the cover image is run through a median filter. Thereafter, the RCNN preserves the original underlying image and creates the connection among the watermark and cover pictures. Comparatively to the approaches mentioned in reference 71, the one under discussion shows improved robustness and consistency.

4. Challenges and Open Research Directions

Deep-learning methods used in watermarking have demonstrated strong learning ability and provide exact and superior results. Still, the protection of security and privacy for media data as well as deep-learning models is challenging. Deep-learning-based watermarking techniques have come a long way, yet still major difficulties remain. The most important challenges in this given field are succinctly summarized below:

1. **Trade-off Maintenance**: One major challenge is reaching a harmonic balance among the resilience, imperceptibility, and embedding capability. Most watermarking systems find it difficult to simultaneously optimize these parameters, thereby affecting the general operating efficiency.

2. **Data Security and Model Complexity**: Many methods fall short in adequately managing the security of the watermarking data and the complexity of deep-learning models, therefore producing unsolved flaws in model security..

3. **Domain Method Constraints**: Transform area-based watermarking approaches show better robustness than spatial area methods according empirical data. Consequently, it is essential to solve the limitations of reliance on a single domain based technique in deep-learning based watermarking.

4. **Encryption and Complexity**: Including watermarking and encryption will greatly increase security. Still, this approach accentuates the complexity of the system, which makes its use challenging with efficiency..

5. **Pre-trained Models**: Pre-trained models are often employed conventionally to simplify model training, therefore reducing the complexity and leading to issues like model overwriting and susceptibility to surrogate model attacks..

6. **Training Dependence**: The choice of loss function and the quantity of training samples define the robustness of watermarking significantly. Either insufficient training data or improper use of the loss function may reduce the watermark's effectiveness..

7. **Poor Embedding Capacity**: Most deep-learning based watermarking techniques have low embedding capacity, which limits their practical relevance.

8. **Robustness Against Pruning and Fine-tuning**: Watermarking techniques should be strong enough to resist minor fine-tuning and network pruning. One still major yet unsolved problem is assurance of this resiliency..

9. **Complete Security**: The continuous challenge of guaranteeing best security for digital content nevertheless remains unanswered. More research is needed to create deep learning-based techniques that provide complete security of digital assets within the framework.

5. Conclusion

DL has progressed greatly in the realm of picture dispensation and now affects data hiding to provide consistent and efficient watermarking techniques. The common deep-learning models used in watermarking systems are thoroughly surveyed in this work along with their roles, purposes, and possibilities to overcome conventional restrictions. We first go over conventional watermarking methods, stressing their fundamental drawbacks—that is, difficulties preserving the trade-off between embedding capacity, imperceptibility, and durability. Classical techniques can rely heavily on manual feature engineering, lack flexibility, and have difficulty resisting challenging assaults. By comparison, deep-learning methods give a good answer to these difficulties as they provide improved adaptation to many media, automated feature extraction, and more resilience. Emphasizing the necessity to address security, model complexity, and performance in real-world applications, this work then describes the fundamental prerequisites for building DL-based watermarking.

We investigate the functions and uses of several deep-learning models in watermarking: autoencoders, convolutional neural networks (CNNs), generative adversarial networks (GANs), and recurrent neural networks (RNNs). We investigate in watermark embedding and extraction the particular goals, techniques, and purposes for every model. We also investigate how they affect things such computing cost, visual quality, resilience against assaults, and ideal embedding sites. Furthermore included in this study are the findings, constraints, and current advancements of many deep-learning watermarking techniques in improving copyright protection, media authentication, and model security.

Deep-learning-based watermarking methods suffer various problems even with their promise. These include preserving an ideal balance between resilience, imperceptibility, and capacity; handling the complexity brought forth by encryption integration; and minimizing vulnerabilities like surrogate model assaults and model overwriting. Furthermore, even while transform domain-based approaches usually show more resilience than spatial domain methods, building solutions combining many domain techniques is still a subject of active investigation. We also discuss the challenges of boosting embedding capacity, assuring model training with enough sample sizes to lower errors, and enhancing watermark resistance against network pruning and fine-tuning. This work aims to provide some insight on the use of deep-learning models in watermarking techniques, therefore acting as a helpful tool for field researchers. Presenting the strengths, uses, and current constraints of deep-learning-based watermarking, we aim to motivate future research into developing more safe, strong, and useful watermarking systems. Although deep learning has great promise in this field, our study emphasizes the need of continuous research to solve the remaining problems and improve the efficiency of watermarking techniques.

References

[1] Singh, A. K. (2019). Robust and distortion control dual watermarking in LWT domain using DCT and error correction code for color medical image. Multimedia Tools and Applications, 78(21), 30523-30533.

[2] Singh, O. P., Singh, A. K., Srivastava, G., & Kumar, N. (2021). Image watermarking using soft computing techniques: A comprehensive survey. Multimedia Tools and Applications, 80(20), 30367-30398.

[3] Singh, O. P., Singh, A. K., Agrawal, A. K., & Zhou, H. (2022). SecDH: Security of COVID-19 images based on data hiding with PCA. Computer Communications, 191, 368-377.

[4] Singh, H. K., & Singh, A. K. (2023). Comprehensive review of watermarking techniques in deep-learning environments. Journal of Electronic Imaging, 32(3), 031804-031804.

[5] Gull, S., Loan, N. A., Parah, S. A., Sheikh, J. A., & Bhat, G. M. (2020). An efficient watermarking technique for tamper detection and localization of medical images. Journal of ambient intelligence and humanized computing, 11, 1799-1808.

[6] Singh, D., & Singh, S. K. (2017). DCT based efficient fragile watermarking scheme for image authentication and restoration. Multimedia Tools and Applications, 76, 953-977.

[7] Huang, Z., Lin, Y., & Chen, X. (2024). A block-based adaptive high fidelity reversible data hiding scheme in interpolation domain. Multimedia Tools and Applications, 83(22), 61715-61736.

[8] Kamaruddin, N. S., Kamsin, A., Por, L. Y., & Rahman, H. (2018). A review of text watermarking: theory, methods, and applications. IEEE Access, 6, 8011-8028.
[9] Amrit, P., & Singh, A. K. (2022). Survey on watermarking methods in the artificial intelligence domain and beyond. Computer Communications, 188, 52-65.
[10] Li, Y., Wang, H., & Barni, M. (2021). A survey of deep neural network watermarking techniques. Neurocomputing, 461, 171-193.
[11] Wan, W., Wang, J., Zhang, Y., Li, J., Yu, H., & Sun, J. (2022). A comprehensive survey on robust image watermarking. Neurocomputing, 488, 226-247.
[12] Zhang, C., Lin, C., Benz, P., Chen, K., Zhang, W., & Kweon, I. S. (2021). A brief survey on deep learning based data hiding. arXiv preprint arXiv:2103.01607.
[13] Y Li, Y., Wang, H., & Barni, M. (2021). A survey of deep neural network watermarking techniques. Neurocomputing, 461, 171-193.
[14] Wang, Z., Byrnes, O., Wang, H., Sun, R., Ma, C., Chen, H., .. & Xue, M. (2023). Data hiding with deep learning: A survey unifying digital watermarking and steganography. IEEE Transactions on Computational Social Systems, 10(6), 2985-2999.
[15] Heaton, J. (2018). Ian goodfellow, yoshua bengio, and aaron courville: Deep learning: The mit press, 2016, 800 pp, isbn: 0262035618. Genetic programming and evolvable machines, 19(1), 305-307.
[16] Hatoum, M. W., Couchot, J. F., Couturier, R., & Darazi, R. (2021). Using deep learning for image watermarking attack. Signal Processing: Image Communication, 90, 116019.
[17] Ingaleshwar, S., & Dharwadkar, N. V. (2023). Water chaotic fruit fly optimization-based deep convolutional neural network for image watermarking using wavelet transform. Multimedia Tools and Applications, 82(14), 21957-21981.
[18] Bagheri, M., Mohrekesh, M., Karimi, N., & Samavi, S. (2020). Adaptive control of embedding strength in image watermarking using neural networks. arXiv preprint arXiv:2001.03251.
[19] Taye, M. M. (2023). Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. Computation, 11(3), 52.
[20] Rouhani, B. D., Chen, H., & Koushanfar, F. (2018). Deepsigns: A generic watermarking framework for ip protection of deep learning models. arXiv preprint arXiv:1804.00750.
[21] Sharma, V. K., & Mir, R. N. (2022). An enhanced time efficient technique for image watermarking using ant colony optimization and light gradient boosting algorithm. Journal of King Saud University-Computer and Information Sciences, 34(3), 615-626.
[22] Zhang, L., & Wei, D. (2020). Image watermarking based on matrix decomposition and gyrator transform in invariant integer wavelet domain. Signal Processing, 169, 107421.
[23] Zheng, W., Mo, S., Jin, X., Qu, Y., Deng, F., Shuai, J., .. & Long, S. (2018, May). Robust and high capacity watermarking for image based on DWT-SVD and CNN. In 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA) (pp. 1233-1237). IEEE.
[24] Li, D., Deng, L., Gupta, B. B., Wang, H., & Choi, C. (2019). A novel CNN based security guaranteed image watermarking generation scenario for smart city applications. Information Sciences, 479, 432-447.
[25] Nie, H., Lu, S., Wu, J., & Zhu, J. (2024). Deep Model Intellectual Property Protection with Compression-Resistant Model Watermarking. IEEE Transactions on Artificial Intelligence.
[26] Etemad, E., Samavi, S., Reza Soroushmehr, S. M., Karimi, N., Etemad, M., Shirani, S., & Najarian, K. (2018). Robust image watermarking scheme using bit-plane of hadamard coefficients. Multimedia Tools and Applications, 77, 2033-2055.
[27] Su, Q., & Chen, B. (2018). Robust color image watermarking technique in the spatial domain. Soft Computing, 22, 91-106.
[28] Zizhuo, W. A. N. G., Kun, H. U., HUANG, C., Zixuan, H. U., Shuo, Y. A. N. G., & Xingjun, W. A. N. G. (2024). Robust blind image watermarking based on interest points. Virtual Reality & Intelligent Hardware, 6(4), 308-322.
[29] Plata, M., & Syga, P. (2020, December). Robust spatial-spread deep neural image watermarking. In 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom) (pp. 62-70). IEEE.
[30] Zhu, J. (2018). HiDDeN: hiding data with deep networks. arXiv preprint arXiv:1807.09937.
[31] Luo, X., Zhan, R., Chang, H., Yang, F., & Milanfar, P. (2020). Distortion agnostic deep watermarking. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 13548-13557).
[32] Ahmadi, M., Norouzi, A., Karimi, N., Samavi, S., & Emami, A. (2020). ReDMark: Framework for residual diffusion watermarking based on deep networks. Expert Systems with Applications, 146, 113157.
[33] Mun, S. M., Nam, S. H., Jang, H. U., Kim, D., & Lee, H. K. (2017). A robust blind watermarking using convolutional neural network. arXiv preprint arXiv:1704.03248.
[34] Singh, O. P., & Singh, A. K. (2023). Data hiding in encryption–compression domain. Complex & Intelligent Systems, 9(3), 2759-2772.
[35] Singh, A. K., Dave, M., & Mohan, A. (2016). Hybrid technique for robust and imperceptible multiple watermarking using medical images. Multimedia Tools and Applications, 75, 8381-8401.
[36] Thakur, S., Singh, A. K., Kumar, B., & Ghrera, S. P. (2020). Improved DWT-SVD-based medical image watermarking through hamming code and chaotic encryption. In Advances in VLSI, Communication, and Signal Processing: Select Proceedings of VCAS 2018 (pp. 897-905). Springer Singapore.
[37] Anand, A., & Singh, A. K. (2020). An improved DWT-SVD domain watermarking for medical information security. Computer Communications, 152, 72-80.
[38] Anand, A., Singh, A. K., Lv, Z., & Bhatnagar, G. (2020). Compression-then-encryption-based secure watermarking technique for smart healthcare system. IEEE MultiMedia, 27(4), 133-143.
[39] Hsu, L. Y., & Hu, H. T. (2021). QDCT-based blind color image watermarking with aid of GWO and DnCNN for performance improvement. IEEE Access, 9, 155138-155152.
[40] Byun, S. W., Son, H. S., & Lee, S. P. (2019). Fast and robust watermarking method based on DCT specific location. IEEE Access, 7, 100706-100718.
[41] Li, J., Lin, Q., Yu, C., Ren, X., & Li, P. (2018). A QDCT-and SVD-based color image watermarking scheme using an optimized encrypted binary computer-generated hologram. Soft Computing, 22, 47-65.
[42] Chen, B., Zhou, C., Jeon, B., Zheng, Y., & Wang, J. (2018). Quaternion discrete fractional random transform for color image adaptive watermarking. Multimedia Tools and Applications, 77, 20809-20837.
[43] Moosazadeh, M., & Ekbatanifard, G. (2019). A new DCT-based robust image watermarking method using teaching-learning-based optimization. Journal of Information Security and Applications, 47, 28-38.
[44] Kandi, H., Mishra, D., & Gorthi, S. R. S. (2017). Exploring the learning capabilities of convolutional neural networks for robust image watermarking. Computers & Security, 65, 247-268.
[45] Nagai, Y., Uchida, Y., Sakazawa, S., & Satoh, S. I. (2018). Digital watermarking for deep neural networks. International Journal of Multimedia Information Retrieval, 7, 3-16.
[46] Uchida, Y., Nagai, Y., Sakazawa, S., & Satoh, S. I. (2017, June). Embedding watermarks into deep neural networks. In Proceedings of the 2017 ACM on international conference on multimedia retrieval (pp. 269-277).

[47] Guan, X., Feng, H., Zhang, W., Zhou, H., Zhang, J., & Yu, N. (2020, October). Reversible watermarking in deep convolutional neural networks for integrity authentication. In Proceedings of the 28th ACM International Conference on Multimedia (pp. 2273-2280).
[48] Wu, J., Shi, H., Zhang, S., Lei, Z., Yang, Y., & Li, S. Z. (2018, February). De-Mark GAN: Removing dense watermark with generative adversarial network. In 2018 International Conference on Biometrics (ICB) (pp. 69-74). IEEE.
[49] Cu, V. L., Burie, J. C., Ogier, J. M., & Liu, C. L. (2019, September). A robust data hiding scheme using generated content for securing genuine documents. In 2019 International Conference on Document Analysis and Recognition (ICDAR) (pp. 787-792). IEEE.
[50] Loc, C. V., Burie, J. C., & Ogier, J. M. (2018, April). Stable regions and object fill-based approach for document images watermarking. In 2018 13th IAPR International Workshop on Document Analysis Systems (DAS) (pp. 181-186). IEEE.
[51] Loc, C. V., Burie, J. C., & Ogier, J. M. (2018, August). Document images watermarking for security issue using fully convolutional networks. In 2018 24th International conference on pattern recognition (ICPR) (pp. 1091-1096). IEEE.
[52] Cu, V. L., Burie, J. C., & Ogier, J. M. (2018, August). Watermarking for security issue of handwritten documents with fully convolutional networks. In 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR) (pp. 303-308). IEEE.
[53] Wei, Q., Wang, H., & Zhang, G. (2020). A robust image watermarking approach using cycle variational autoencoder. Security and Communication Networks, 2020(1), 8869096.
[54] Zhang, L., Li, W., & Ye, H. (2021, October). A blind watermarking system based on deep learning model. In 2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom) (pp. 1208-1213). IEEE.
[55] Fan, B., Li, Z., & Gao, J. (2022). DwiMark: a multiscale robust deep watermarking framework for diffusion-weighted imaging images. Multimedia Systems, 28(1), 295-310.
[56] Fang, H., Jia, Z., Zhou, H., Ma, Z., & Zhang, W. (2022). Encoded feature enhancement in watermarking network for distortion in real scenes. IEEE Transactions on Multimedia, 25, 2648-2660.
[57] Kang, X., Huang, J., & Zeng, W. (2010). Efficient general print-scanning resilient data hiding based on uniform log-polar mapping. IEEE Transactions on Information Forensics and Security, 5(1), 1-12.
[58] Zhu, J. (2018). HiDDeN: hiding data with deep networks. arXiv preprint arXiv:1807.09937.
[59] Liu, Y., Guo, M., Zhang, J., Zhu, Y., & Xie, X. (2019, October). A novel two-stage separable deep learning framework for practical blind watermarking. In Proceedings of the 27th ACM International conference on multimedia (pp. 1509-1517).
[60] Ma, Z., Zhang, W., Fang, H., Dong, X., Geng, L., & Yu, N. (2021). Local geometric distortions resilient watermarking scheme based on symmetry. IEEE Transactions on Circuits and Systems for Video Technology, 31(12), 4826-4839.
[61] Neekhara, P., Hussain, S., Zhang, X., Huang, K., McAuley, J., & Koushanfar, F. (2022). FaceSigns: semi-fragile neural watermarks for media authentication and countering deepfakes. arXiv preprint arXiv:2204.01960.
[62] Hou, J., Ou, B., Tian, H., & Qin, Z. (2021). Reversible data hiding based on multiple histograms modification and deep neural networks. Signal Processing: Image Communication, 92, 116118.
[63] He, W., Cai, J., Zhou, K., & Xiong, G. (2017). Efficient PVO-based reversible data hiding using multistage blocking and prediction accuracy matrix. Journal of Visual Communication and Image Representation, 46, 58-69.
[64] Jia, Y., Yin, Z., Zhang, X., & Luo, Y. (2019). Reversible data hiding based on reducing invalid shifting of pixels in histogram shifting. Signal Processing, 163, 238-246.
[65] Li, X., Zhang, W., Gui, X., & Yang, B. (2015). Efficient reversible data hiding based on multiple histograms modification. IEEE Transactions on Information Forensics and Security, 10(9), 2016-2027.
[66] Zhang, J., Gu, Z., Jang, J., Wu, H., Stoecklin, M. P., Huang, H., & Molloy, I. (2018, May). Protecting intellectual property of deep neural networks with watermarking. In Proceedings of the 2018 on Asia conference on computer and communications security (pp. 159-172).
[67] Wu, H., Liu, G., Yao, Y., & Zhang, X. (2020). Watermarking neural networks with watermarked images. IEEE Transactions on Circuits and Systems for Video Technology, 31(7), 2591-2601.
[68] Deeba, F., Kun, S., Dharejo, F. A., Langah, H., & Memon, H. (2020). Digital watermarking using deep neural network. International Journal of Machine Learning and Computing, 10(2), 277-282.
[69] Hamamoto, I., & Kawamura, M. (2019). Image watermarking technique using embedder and extractor neural networks. IEICE transactions on Information and Systems, 102(1), 19-30.
[70] Sinhal, R., Jain, D. K., & Ansari, I. A. (2021). Machine learning based blind color image watermarking scheme for copyright protection. Pattern Recognition Letters, 145, 171-177.
[71] Islam, M., Roy, A., & Laskar, R. H. (2018). Neural network based robust image watermarking technique in LWT domain. Journal of Intelligent & Fuzzy Systems, 34(3), 1691-1700.
[72] Kazemi, M. F., Pourmina, M. A., & Mazinan, A. H. (2020). Analysis of watermarking framework for color image through a neural network-based approach. Complex & Intelligent Systems, 6, 213-220.
Singh, A. K., Kumar, B., Singh, S. K., Ghrera, S. P., & Mohan, A. (2018). Multiple watermarking technique for securing online social network contents using back propagation neural network. Future Generation Computer Systems.