Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

An Overview of The Proposed Technique for Removing Visible Watermark

Ali Abdulazeez Mohammedbaqer Qazzaz, Zahraa Mohammed Jabbar,

University of Kufa / Faculty of Education / Department of Computer
*Corresponding author E-mail: zahraam.altufaili@student.uokufa.edu.iq

https://doi.org/10.46649/fjiece.v3.2.36a.15.6.2024

Abstract. *Watermarks are commonly used to protect the ownership and copyright of digital media. However, there are legitimate scenarios where watermark removal is necessary. Recent advancements in deep learning have led to the development of sophisticated techniques for both detecting and removing watermarks. this research provides a summary of methods for detecting and removing Regenerative Adversarial Networks (GANs) are one noteworthy method. It is possible to train GANs to recognize watermark patterns and produce unwatermarked versions of watermarked content. One such method, which uses GANs to find and remove watermarks in deep neural networks (DNNs), has been demonstrated to be successful even when it comes to DNN watermarks that are based on backdoors. Another method makes use of deep neural networks' U-structure, which is highly effective in translating images. A comprehensive model like the AdvancedUnet has been developed to concurrently extract and remove visual watermarks. This model uses a deep-supervised hybrid loss to direct the network in learning the transformation between the watermarked input and the clean ground truth. It also integrates efficient modules to extend the architecture's depth without appreciably increasing computing costs.*

Keywords: deep learning, digital watermarking, GANs, CNNs, DNNs, image security, copy-right protection, machine learning, signal processing

1. INTRODUCTION

In recent years, deep learning has advanced different instance operations including classification, segmentation, super-resolution, deblurring, and denoising to create images [1]. This technique employ machine learning algorithm trained to construct representations of data directly from raw images, thereby bypassing the need for manual feature engineering. Diagnostic images are extensively employed in areas like medicine, communications, forensics, education and the research and development. Images data, especially that of secret personal or the organizational matter could not be disclosed to outsiders without permission. Security of digital data and avoiding copyright infringement through watermarking techniques are researched by cyber researchers. Such a method will ease up the transmission of mobile phones in smart devices on unsecured channels and preserve the original data mostly in the process [2].

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Figure 1: Recent applications of watermarking

A conventional watermarking technique embeds confidentiality key or authorship information into an image and then disseminates it across the public networks. This digital signature confirms that an image is legitimate and hasn't been manipulated. The basic structure of the watermarking process, which primarily consists of two stages: throw and take out [3].

The watermark, cover media and the secret key are fed to the entrust algorithm that overwrites the watermark into the cover media via this mechanism generating watermarked content. Encoding a block of k-bits in the domains of example - either spatial or transform domain methods is done by algorithm. The watermarking process is also more secure than traditional copyright registration method since it involves caption, encryption, encoding, and haschement among others as noted in [4, 5, 6].

The focus on deep-learning-based watermarking techniques in the recent years has brought great successes and the results are found to be astoundingly great compared to the conventional methods [1, 8, 9]. When employing deep learning for watermarking, critical advantages include [1, 10]: (a) construction of tough watermarks, (b) finding the right points for tracks embedding in secret media, (c) selecting the highest strength embedding power to make in some way a balance between quality and robustness, (d) using simulations of attacks to improve watermarks extraction, (e) improvement of the error suppressing and denoising of the recovered watermarks. But, data security and privacy are other major issues faced by deep learning techniques (as discussed by 13) as well. Today, the deep-learning-based watermarking technique in particular, provides brands a powerful tool to prevent misappropriations and deal with the issue of arguing over copyright infringement in terms of ownership, in which watermark provides a platform on which the identity of originality has a role to play for verifying the ownership.

Several questionnaires have been published in the past few years about the protection of media and ground zero models by the means of watermarking intelligence [1, 8, 11, 12]. Research Article [1] summarizes about deep learning model usage in the watermarking process which includes the different stages. Instructiion [8] discusses various techniques in watermarking applied on the artificial intelligence basis. Reference [11] discusses the issue of protecting of the deep learning models by means of watermarking and Reference [12] contributes with the security mechanisms for the intellectual property of deep learning techniques. Rather, our input focus is the location of the disjunction which is Digital Watermarking employment of the prominent aspects of deep learning models that includes copyright and ownership traces and also securing the models. We go on to consider the role and effect of some deep-learning models on watermarking by indicating that they are we cases in the studies analyzed.

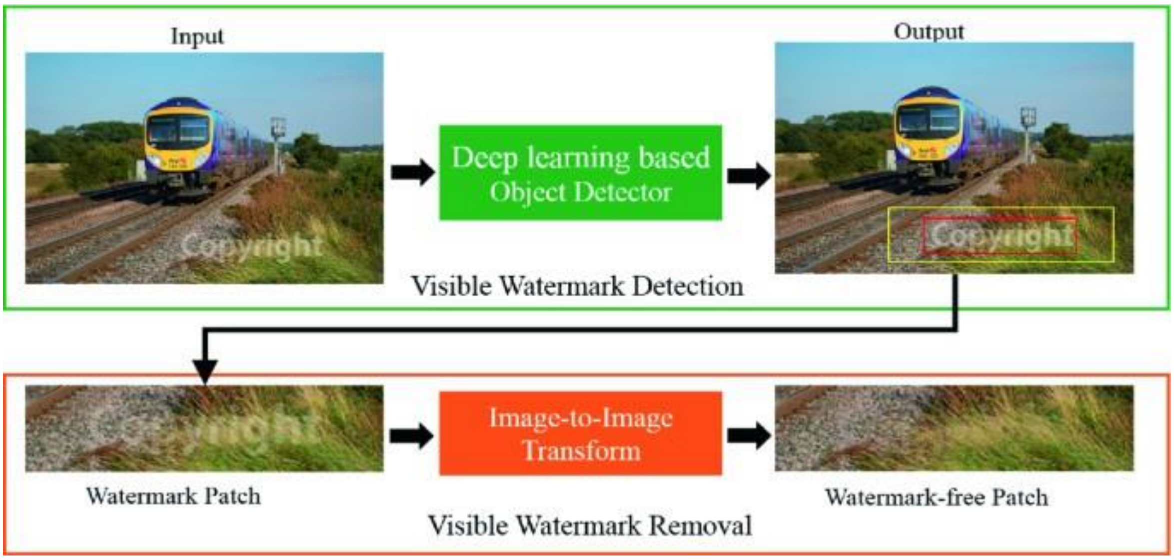


Figure 2: The framework locates, magnifies, and creates input for the purpose of removing watermarks from images.[13]

Figure 2 illustrates the functioning of a vision-based object detection system based on deep learning. This neural network completes the task of image processing, determining and emphasizing the elements of "copyright" watermarks present in a photo with a train sitting on the tracks against a rural background. The next step is a close-up inspection of the watermark in its own right. The zone that is going to be used for the further processing and referred to as the "Watermark Patch," is selected on the picture. The patch holds the vector watermark that will pass through the "Image to Image conversion" step soon.

This last step is the deep learning model based operation that removes watermark markings. The generated outcome, as the image is presented, is a "Watermark-free Patch," a free section of the image that such a watermark has been totally removed. The removal process, however, is made particularly gentle which makes it possible to obtain an image that is barely imperceptible from the original. The final outcome is comprised of the original photo plus it is watermark-free; therefore it can be used at leisure without the copyright Legend.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

2. Watermark Detection

The researchers in [14] devised a technique which case of pictures to be watermarked, we may use a Mask R-CNN to get the embedding strength. Both DCT and DWT regions were selected at first and processed by using lower level ROI pixels. After, the watermark was put inside the cover image. While basically successful watermark extraction was recorded this method cannot provide a broad framework for embedding and watermark security. A wavelet-based strategy to evaluate invisibility and robustness when watermark is added to the image was illustrated by Zheng et al. [15]. Typically, extraction is merely the reversal of embedding, which may be formulated in a simple way as follows: We will use M , V , and S to denote the covering medium, watermark, and super secret key, respectively, and the embedding function will be abbreviated as $Embed()$. The watermarked media M' is mathematically expressed as follows:We represent watermarked media $M'+$ in the following mathematical way:

$$M' \approx (1 - w) * M + w * N \qquad (1)$$

where the term "equation" is denoted as Eq. (1). The algorithm can make use of multiple different techniques such as the pictures-based, phrase-based and structure-based [7] method to encrypt the watermark of the crisis.

The watermarked media M' which was passed through an open and insecure network has more chances of deformation or with kind of misbehaviour. It can be torpedoed enduringly(deliberately by attackers), or accidentally (by noise), calling hence proving that the watermark can be so well readable in order to cater the copyright protection and at the same time the information can prove to be confidential post the attack.

For Extract$(\cdot, \cdot)$ is the extraction function and M'' depict the media received after watermarked. The extraction of the watermark W' is mathematically defined as: The extraction of the watermark W' is mathematically defined as:

$$(W' \rightarrow (M'')\text{Extract} \qquad (2)$$

in this case anyone will agree that Eq (2).

In the beginning, the frequency range of the cover media was separated into different frequency bands (via DWT's partition), and data singular values (from the watermark) was added to the upper band. The watermark sequence was used thereafter embedded in the bands that were designated as low bands through wavelet transformation. The block of ice, cover media, and the watermarked media were related employing CNN to promote strong and true watermark extraction. These characters bear the cost that is below their rivals', and this makes it more useful in practice. The watermark is embedded into a cover image having undergone the DWT treatment by means of a CNN method presented in [16] and consisting of a DWT with a scrambled output. The networks apply a fast region-based CNN model to demonstrate more robust and blind extraction of the watermarks in the embedded warping. Its offer in this regard includes high classification accuracy, invisibility and less dependence on execution time when compared to classical techniques. . However, in addition to a comprehensive elucidation, real-world testing is also needed to ascertain many aspects related

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

to images' manipulations resilience against different processing attacks. In [17], they proposed a watermarking method based on a CNN (Convolutional Neural Network) model whereas a spatial watermarking information is embedded into the cover picture. They devised (came up with) a loss function while training the neural network to improve the robustness and at the same time keep a balance other tradeoffs of watermarking. It has been also assessed that this method shows high robustness at the cost of marginal distortion and if we accept it more studies should be undertaken in order to evaluate its performance under other image processing attacks. Instead, Mun and his co-authors [18] as a watershed in the network for blind watermarking based on deep CNN images to control the copyright, it was suggested to break the carrier and media into non-overlapping units. To implement watermark embedding, the deep CNN model was utilized.Finally, the model was pulled out for the information gathering purposes from the book cover. These attacks, therefore, are characterized by the necessity of choosing between robustness, anti-discriminability and cost simultaneously because of their overall effectiveness in responding to different Salt and pepper noise attacks. Table 1 Summary of CNN-based Watermarking Techniques Employed for Enhancing Image Security and Robustness in Digital Media. In contrast to methodologies stated in [19, 20, 21, 22], the classical defense showed 27.84% improvement in damage resistance.

Table 1: Summary of CNN-based Watermarking Techniques Employed for Enhancing Image Security and Robustness in Digital Media.

Reference	**Method**	**Key Features**	**PSNR**	**Robustness**
[23]	Optimization-based deep CNN	Grid feature extraction and optimization technique for embedding	High	Moderate
[14]	Blind watermarking with Mask R-CNN	Uses lower ROI pixels in DCT and DWT blocks for embedding	Moderate	Limited
[15]	Wavelet-based watermarking	Inserts watermark in high and low DWT bands, uses CNN for extraction	Good	High
[16]	Fast region-based CNN	Embeds scrambled watermark into DWT cover image for robust extraction	Very High	Very High
[17]	Spatial data embedding CNN	Trains with designed loss function for improved trade-offs	Not specified	High
[18]	Blind watermarking with deep CNN	Uses non-overlapping blocks for embedding and extraction	Good	Very High

Reference	Method	Key Features	PSNR	Robustness
[26]	Auto-encoder CNN	Uses CNN auto-encoder capabilities for robust embedding	Not specified	Prone to at-Tacks
[27]	Watermark for CNN ownership	Embeds watermark into convolution layers of CNN	Not specified	Limited to specific attacks
[28]	Pruning-based watermarking in ResNet 152	Embeds SHA-256 hash in CNN layers using pruning theory	Not specified	High
[25]	QDCT-based method	Uses grey wolf optimizer for robustness and imperceptibility balance	Not specified	High
[24]	Encryption-compression technique	LWT, RSVD, HD, and CNN-based denoising for better extraction	Improved by 27.84%	High

In watermarking, the use of Generative Adversarial Network (GAN) deep learning model has been popular due to its dual-network structure comprising of a generator and a discriminator. These network are powerful in generating and verifying watermarks. GANs are specifically used for network security by attaching to each input is watermark of unique type. The generator network G takes a random noise input coming from distribution $P(Z)$ and then maps it to the sample S where probability distribution of $P(X)$ is mimicked. Whereas, the discriminator network (D) determines the artificial models from the authentic ones. As it goes under the training process, the discriminator learns to use features that can protect the model from wrong rejection of true or vice versa. The loss function L used in GANs is defined by the equation: The loss function L used in GANs is defined by the equation:

This is equivalent to finding the minimizer of the surrogate function which yields its maximum value of

$$L = \min \max [\log(D(x)) + \log(1 - D(G(z)))]$$

They called their model a discriminator generated by deep learning, which closely emulated the distention, and their image resynthesis was tapped to increase quality and promise of success. The generator that is an autoencoder, performing these operations for the artificially damaged pictures into a representation vector, however, the instruction module complete this decoding into RGB pictures. The source description guides the generator output as it minimizes the feature loss in turn this leading generation of the true images of the founder. Here, his aim is to utilize a multi-layer convolutional network

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

(ResNet-46) which with the fila recovered image will generate features as well as it will generate the original one. With verification accuracy of 96.36% and False Positive Rate (FPR) of 1%, this model had a very useful progress and final result regarding the prevention of food fraud.

Wei et al. [29] worked out a kind of watermarking strategy combining the variational autoencoder network with the copyright protected bottom line. The system of subnetworks representing the encoder, decoder, and detector is such type. In training a 1 bit watermark image is embedded into your host image while encoder and decoder subnetworks develop a robust representation to a cover image being translated.

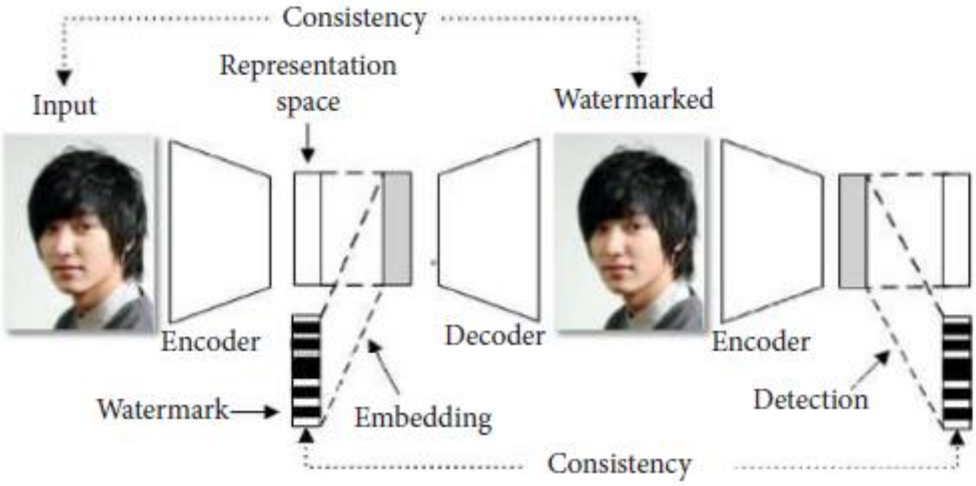


Figure 3: *Framework of Cycle-VAE model for image watermarking.* [29]

It is the job of the detector subnetwork to learn to capture and preserve the 1-bit evidences embedded in the watermarked image. It is true though that the image quality of the watermarked picture has improved but for the time being, there is still a need for further investigations to ensure the reliability of the watermarked images under different situations. Deep learning was applied as a watermarking approach, which completely did not detect [30]. This method includes four components: neural network consisting of an encoder, decoder, two channels of noise, and one adversarial discriminator. Each of the layers carrying the same mark performs a watermark embedding and extraction process using the encoder-decoder pair thus, making a watermark highly robust to different types of attacks. Moreover, an adversarial discriminator helps in enhancing the robustness and hideousness of the water mark too. In terms of general performance, the architecture is certainly an effective tool against many types of incidents; it has, however, a very high complexity.

Fan’s et al.’s approach [31] utilizes multiscale robust watermarking technology to prevent diffusion-weighted imaging (DWI) images from being modified through the process of transmission or being accessed by unauthorized parties. This technique uses a multiscale strategy and adversarial network of a generative nature. Initially, UDI data is analysed and extracted into values that are matched with the original images. Subsequently, watermarks are hidden into multiscale feature series of the reconstructed images. A well configured boundary equilibrium adversarial network generative discriminator is proposed to obtain the pixel quality benchmark image with the pyramid filters and multiscale max-pooling to learn the texture distribution map.

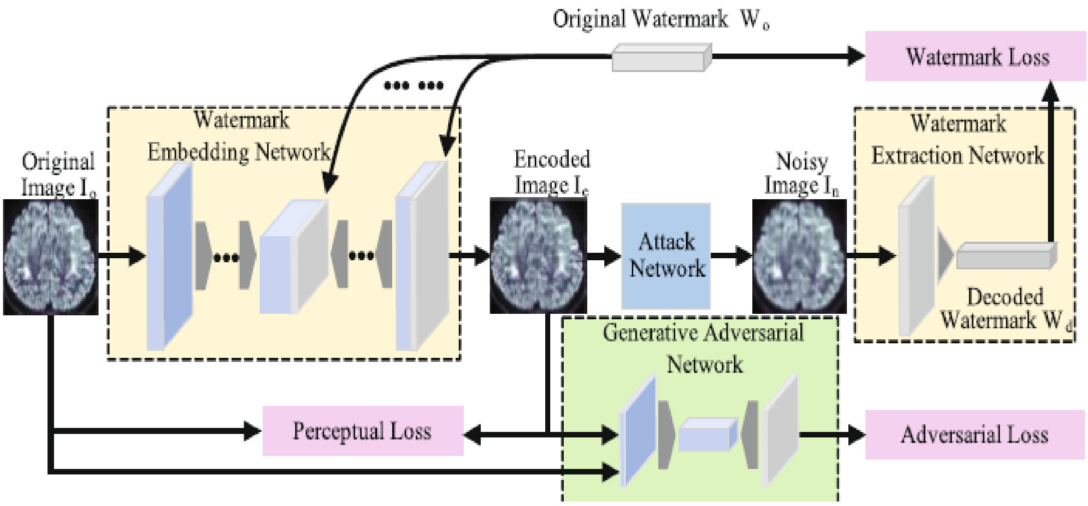
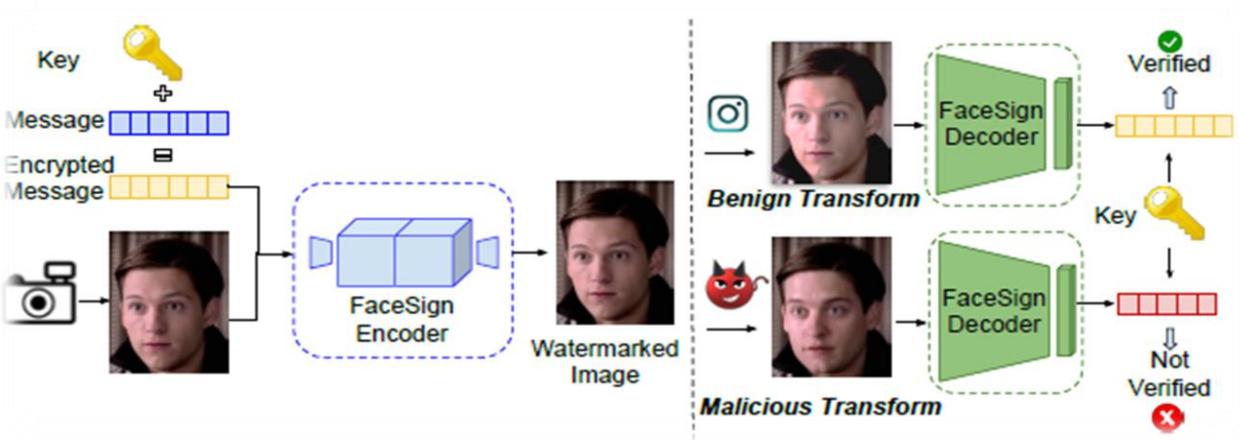


Figure 4: There are four steps involved in the embedding and extraction of watermarks[31]

Fang et .al [32] proposed a unique watermarking scheme comprising triple phase to keep the image clear in operation. This strategy applies an unbiased phase, a frequency augmentation phase, and a adversarl phase. In the beginning, an encoder–decoder is fine-tuned using just-noticeable difference (JND) mask image loss – a difference that the person who is viewing either cannot notice or only notices with a slight difficulty. After this, there is a feature encoding. Then, a mask-guided frequency method is used for frequency augmentation. Adversarial training is applied to an encoder in the last stage, which allows it to overcome the distortion that is induced by quantization making it more robust than recent works [33, 34, 35, 36].An advanced watermarking algorithm having semi-fragile property based on deep learning for media authentication was put forth in [37]. Table 2 summarizes the most important techniques for removing the watermark using GAN and Table 3 summarizes watermark detection techniques.



Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Figure 5: The FaceSigns Watermarking framework incorporates encrypted text into images[37].

Table 2: Summary of GAN-Based Watermarking Techniques

Ref	Methodology	Key Features	Performance Metrics	Comments on Robustness
[38]	Deep-learning model to verify corrupted images	Generator-discriminator architecture; uses ResNet-46 for feature extraction	96.36% Accuracy at 1%FPR	High verification accuracy
[39]	Robust data hiding using GAN	Geometric correction and adversarial network for document security	Better than references [40, 41, 42]	Improved robustness
[29]	Variational autoencoder for copyright protection	Encoder, decoder, and detector subnetworks 1-bit watermark embedding	Enhances visual quality of marked image	Needs further robustness testing
[30]	Blind watermarking scheme	Includes encoder, decoder, noise layers, and adversarial discriminator	Good imperceptibility	High complexity; robust against various attacks
[31]	Multiscale robust watermarking for DWI images	Uses full-scale features and a generative adversarial network	Enhances visual quality of a reconstructed image	Robust: includes optimized boundary equilibrium GAN discriminator
[32]	Triple-phase watermarking scheme	Noise-free initial Phase, mask-guided frequency augmentation, adversarial training	More robust than references [33, 34, 35, 36]	Highly robust to non-differentiable distortion
[37]	Semi-fragile watermarking for media authentication	Three modules: encoder, decoder, and adversarial discriminator networks	Provides tamper detection	Requires further investigation against more attacks

Table 3: Summary of Watermark Detection Techniques

Ref	Method	Key Features	Performance Metrics	Comments on Robustness
[23]	Optimization-based deep CNN	Grid feature extraction and optimization technique for embedding	High PSNR	Moderate robustness
[14]	Blind watermarking with Mask R-CNN	Uses lower ROI pixels in DCT and DWT blocks for embedding	Moderate PSNR	Limited robustness
[15]	Wavelet-based watermarking	Inserts watermark in high and low DWT bands, uses CNN for extraction	Good PSNR	High robustness

Table 3 Continued from previous page

Ref	Method	Key Features	Performance Metrics	Comments on Robustness
[15]	Wavelet-based watermarking	Inserts watermark in high and low DWT bands, uses CNN for extraction	Good PSNR	High robustness
[16]	Fast region-based CNN	Embeds scrambled watermark into DWT cover image for robust extraction	Very High PSNR	Very high robustness
[17]	Spatial data embedding CNN	Trains with designed loss function for improved trade-offs	Not specified	High robustness
[18]	Blind watermarking with deep CNN	Uses non-overlapping blocks for embedding and extraction	Good PSNR	Very high robustness
[24]	Encryption-Compression technique	LWT, SVD, HD, and CNN-based denoising for better extraction	Improved damage resistance by 27.84%	High robustness
[25]	QDCT-based method	Uses grey wolf optimizer for robustness and imperceptibility balance	Not specified	High robustness
[26]	Auto-encoder CNN	Uses CNN auto-encoder capabilities for robust embedding	Not specified	Prone to attacks
[27]	Watermark for CNN ownership	Embeds watermark into convolution layers of CNN	Not specified	Limited to specific attacks
[28]	Pruning-based watermarking in ResNet 152	Embeds SHA-256 hash in CNN layers using pruning theory	Not specified	High robustness
[38]	Verification of corrupted images using GAN	Generator-discriminator architecture; uses ResNet-46 for feature extraction	96.36% accuracy at 1% FPR	High verification accuracy
[39]	Robust data hiding using GAN	Geometric correction and adversarial network for document security	Better than references [40,41,42]	Improved robustness
[29]	Variational autoencoder for copyright protection	Encoder, decoder, and detector subnetworks; 1-bit watermark	Enhances visual quality of marked image	Needs further robustness testing

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Table 3 Continued from previous page

Ref	Method	Key Features	Performance Metrics	Comments on Robustness
[30]	Blind watermarking scheme	Includes encoder, decoder, noise layers, and adversarial discriminator	Good imperceptibility	High complexity; Robust against various attacks
[31]	Multiscale robust watermarking for DWI images	Uses full-scale features and a generative adversarial network	Enhances visual quality of reconstructed image	Robust includes optimized boundary equilibrium GAN discriminator
[32]	Triple-phase water-marking scheme	Noise-free initial phase, mask-guided frequency augmen-tation, adversarial training	More robust than references [33, 34, 35, 36]	Highly robust non-differentiable distortion
[37]	Semi-fragile water-marking for media authentication	Three modules: en-coder, decoder, and adversarial discrim-inator networks	Provides tamper detection	Requires further investigation against more attacks
[43]	Enhanced multiple histogram modification	Uses DNNs to produce and select optimal histogram bins for embedding	Higher PSNR than refs [44, 45, 46]	Requires further robustness and cost analysis
[47]	Copyright protection and ownership verification	Enhances previous method [27] with a threat model enabling API access	Good accuracy with minimal overhead	Lacks detailed security and overhead analysis
[48]	Copyright and ownership protection using DNN	Outputs from DNNs used to obtain watermarked images; extraction by specific network	Evaluated for three types of attacks	Limited real-time application study
[49]	Intellectual property protection	Embeds watermark in DNN; verification via specific input-output pairs	Performs well under two attack types	Execution time and robustness need further study
[50]	Ownership verification of multimedia documents	Embeds 1-bit bi-nary watermark in DCT coefficients of blocks	Close resemblance to original, acceptable PSNR	Detailed performance metrics not fully explored

3.Watermark Removal

Watermark embedding has utility in implementing copyright protection mechanisms for digital photographs as a pictorial watermark can be started from some region of the background image. On the other hand, given the fact that intentional removal of visible watermark might provoke more aggressive attacks, the emphasis would be on the strengthening watermarks against collisions that make it immaterial to remove its existence. For example, original methods embraced GAN networks for watermarking tagging avoidance [51] whereas a study came with multi-task techniques to predict watermark-free images at the same time the watermark mask and the watermark pattern [52].

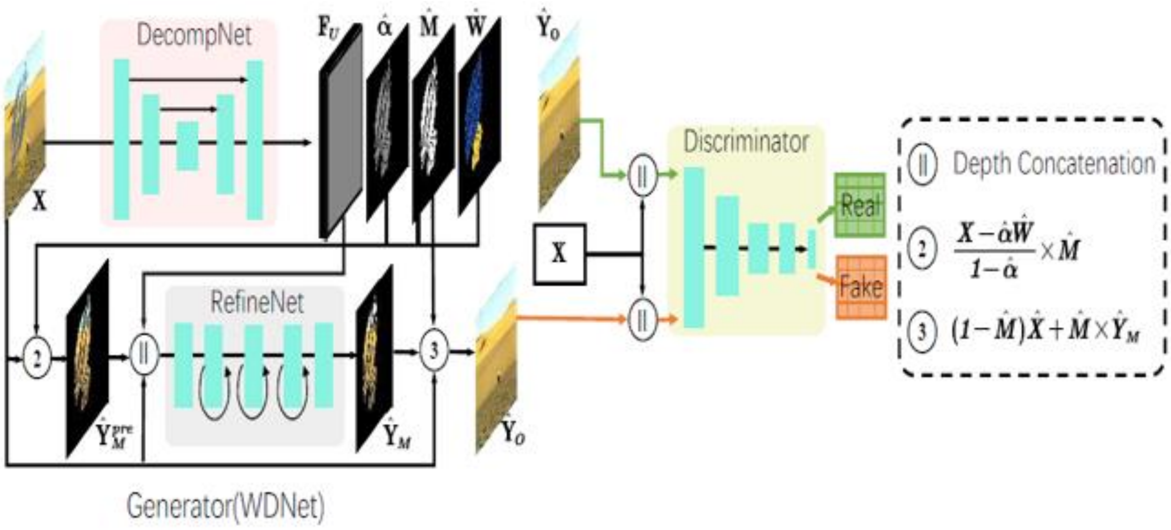


Figure 6: The architecture of our visible watermark removal framework.[52]

Adding to that, the multi-task network that has two stages and also attentive mechanism [53] and another self-calibrated mask refinement to predict the adaptive mask and refine the patient's background [54] were extensively investigated. Semantic similarity-based technique and the dynamic convolution for watermarking focused on versatility was unveiled. The particular themes of the innovations also involve a new architecture for an enhanced long-range information extraction [55], and the intrusion of watermark vaccine that infuses invisible perturbations to fabricate the present watermark decoding techniques [56]. Although there is effort put into this, there are difficulties in identifying complex regions, and many methods still aim at targeting single region watermark without offering the possibility of decomposing a whole watermarked image into various parts. This incorporates a finer-grained version by regulating the amount of local region and parts to be restored adaptively.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

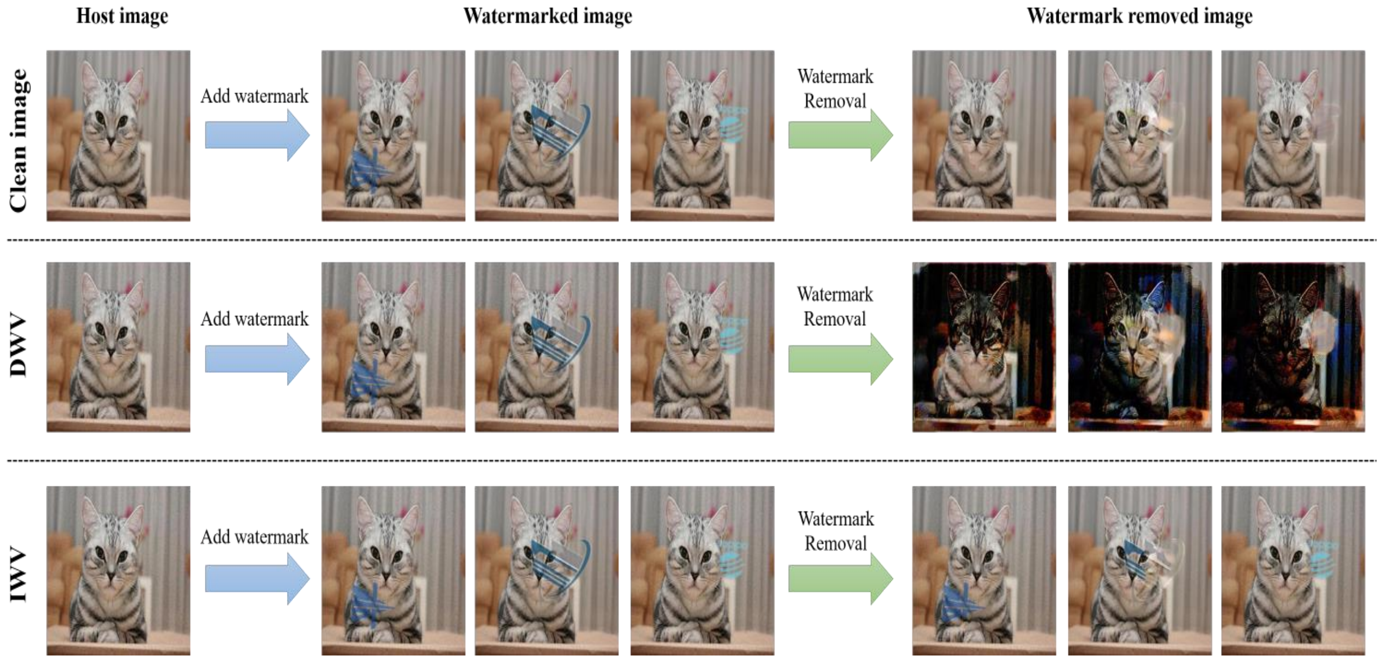


Figure 7: Our watermark vaccines' protective effects on different watermark designs or specifications [56].

Figure 7 shows how our watermark vaccinations protect against various watermark designs or characteristics. The watermarks can be successfully removed using the current blind watermark removal approach, such as WDNet [32] (top). The watermark-removed photos will be destroyed if the host images have the Disrupting Watermark Vaccine (DWV) installed (middle). However, the results cannot be effectively purified as the host images (bottom) when the host photos are loaded with the Inerasable Watermark Vaccine (IWV).

And in the wider background of tasks for image dehazing, image deraining [57] and shadow removal [58], [59, 60] are included. Various universal schemes of image content removal have been suggested like the transformer block with a local window enhancement and a learnable multi-scale restoration modulator [61]. One of the other general blind image decomposition networks is put forward at the same time [62]. One more multi-stage progressive image restoration architecture is suggested as well [63].

Recently, the transformer architecture has made significant inroads into several areas of computer vision, including detection of objects [64], instance and scene segmentation [65], and pose estimation [66]. Also notable are the applications into lower-level tasks such as image super- resolution [67], denoising, and deraining [68], image The network used exploits a query-based multi-task framework relying on adapted query embeddings in both: the mask decoder and the background decoder that, in general terms, resemble other approaches [69] but not taking into account ground-truth part masks.

The networks with dynamically changing parameters numbered among the utilizations of dynamic networks in deep learning image processing tasks date back to adaptive parameters [70], weight predictors [71], and dynamic features [72]. This model uses the adapted query embeddings to support the dynamic control over the weight of the kernel hence display a

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

novel manner of dealing with the different watermarked parts [73]. We will show in Table 4 a summary of watermark removal techniques.

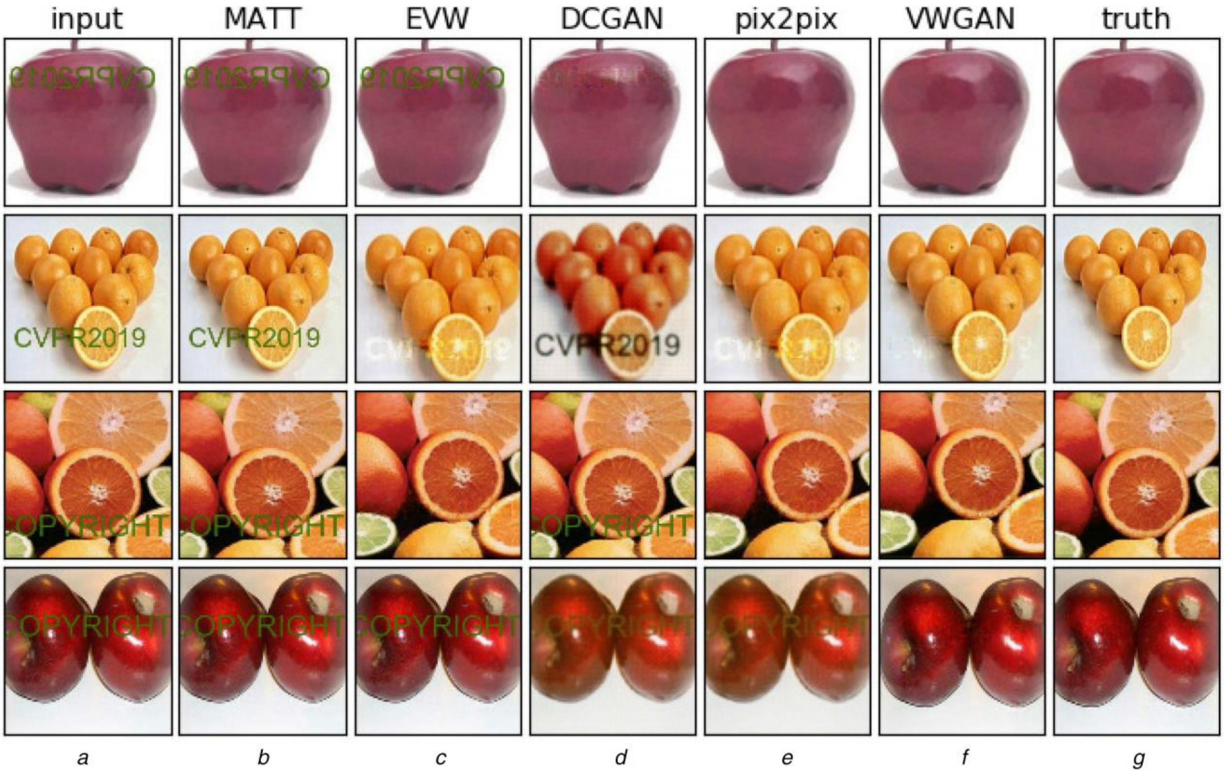


Figure 8: Results of watermarked image removal[73].

Table 4: Summary of Watermark Removal Techniques

Ref	Method	Key Features	Notes
[51], [73]	GAN-based networks	Used for watermark removal	Initial approaches employing GANs to enhance watermark resilience by removing them.
[52], [74]	Multi-task frameworks	Predict watermark-free images, watermark masks, and watermark patterns simultaneously	Introduced to improve the effectiveness of watermark removal through simultaneous processing.
[53]	Two-stage multi-task framework	Incorporates an attention mechanism and a refinement stage	Enhances the precision of watermark removal by focusing on relevant image areas.
[54]	Self-calibrated mask refinement	For adaptive mask prediction and mask-guided background enhancement	Focuses on adapting mask prediction to the specific needs of the background image for better integration.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Table 4: Summary of Watermark Removal Techniques

Ref	Method	Key Features	Notes
[55]	Novel architecture for long-range information extraction	Improved long-range information extraction to deal with complex watermarked regions	Tailored for scenarios where watermark patterns are extensively integrated into the image.
[56]	Watermark vaccine	Introduces invisible perturbations to disrupt watermark removal methods	A defensive approach against unauthorized watermark removal technologies.
[61]	Transformer block with locally enhanced window	Learnable multi-scale restoration modulator	Part of universal methods for broader image con- tent removal challenges, including watermarking.
[62]	General blind image decomposition network	Multi-stage approach for image content handling	Provides a framework that is adaptable for various image restoration tasks including watermark removal.
[63]	Progressive image restoration architecture	Multi-stage architecture designed for detailed and complex image restoration tasks	Suitable for detailed watermark removal as part of broader image restoration efforts.
[75]	Dynamic convolution and semantic similarity	Uses semantic similarity for information propagation to handle diverse watermarks	Developed to address Watermarks dynamically across varied image contexts.

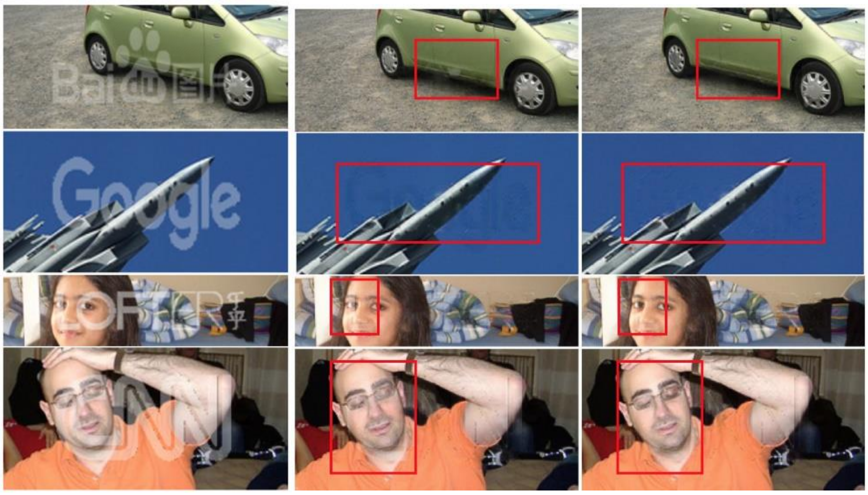


Figure 9: Visible watermarks in real-world situations typically have intricate structures [51]

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

4.Challenges and Future Directions in Deep Learning-Based Watermarking

Through the above of deep learning GAN's CNN's and DNN's approach in digital watermarking along with several challenges arose. Also, some research and development

insights in watermark detection and removal are discussed. GANs being equipped with the generation component, they have been contemplated to bolster watermark safety and reduce the likelihood of an intruder using an existing watermark through generating unique watermark for every input as portrayed in the work of [38]. Yet, there are obstacles to building the watermarking systems that safeguard the copyright-info from a variety of attacks and in its case the watermark which allow for quality maintenance of the content. CNNs are the key things that next to it, there are other methods, just like the one proposed by Ingaleshwar and Dharwadkar [23], have to also envision to solve issues related to computational efficiency and the reliability of the watermark against robust image processing attacks. Among DLs, however, the DNNs are responsible for a great boost in watermarking methods creation, because of their deep architectural structures that show up in process of Hou et al. [43] and Zhang et al. [47] methods development. Despite the fact that the computer networks can be optimized to seemingly cut down the computational burden at the cost of accuracy, the complexity of solving this issue is still clearing majority of the problems[1]. The test of time infrastructure of these networks should be aired out and it takes the prior situation that the hackers will sharpen their methods of removing or spoofing these watermarks[2]. However, outlook prediction can be toward the development of more adaptable and robust deep learning models, which can adjust themselves automatically by recognizing different types of digital media and attack vectors so that the water- marking becomes not so visible for human attacker but at the same time it is unbreakable for any hacking or decoding attacks[3]. On the other hand, incorporation of multi-disciplinary paradigm by using cryptographic knowledge would drive the development of watermarking technology or tools that are more advanced and less invasive[4].

4. CONCLUSIONS

The conclusion of this review is that learning processes, including GANs, CNNs, and DNNs, are becoming an integral part of digital watermarking and many additional solutions related to anti-attacks need to be developed. The rise of these sophisticated neural network models has demonstrated the efficacy of the neural network models in increasing the security of the catalytic watermarking method. GANs, which are unique in digital watermark creation due to their generative capabilities, furnish an array of flawless approaches to generating watermarks that are innately impenetrable by unauthorized entities who use various techniques to detect and alter them. This is because they use their powerful image processing capabilities, which characterize them as representatives for visual tasks demanding total quality of the watermarked content appearance. DNN has real advantages because of its depth and complexity, which can embed and consistently be extracted for watermarks resulting in different conditions .And indeed these approaches are not without formidable barriers. In addition to computing capacity, resistance against strong attacks as well as preservation of data privacy and security integrity are major obstacles to overcome. Additionally, the algorithms that have to be designed to work with different types of media as well as the evolution of attack methods, requires the development of new tools that enhance such processes regularly.

For the future, this matter not only provides an opportunity for deeper investigation but also controversy. The subsequent studies will explore the building of new adaptive and efficient watermarking techniques which are able to be embedded in different types of digital media in a non-invasive manner and can fit in with multiple distribution and media usage scenarios. More- over, multidisciplinary hybrid solutions encompassing parts of cryptography, machine learning, and signal processing are highly prized that they make digital watermarks relevant and acceptable for copyright protection in the digital environment. These technologies will have a bearing on the capability of marking the water by introducing more sustainable, dependable, and user-friendly watermarking options.

REFERENCES

[1] A. A. A. M. B. Tawfiq A. Al-Asadi, "Fusion for Multiple Light Sources in Texture Mapping Object," 2020.
[2] Ali Abdulazeez Mohammed Baqer Qazzaz a, Elaf J. Al Taee a, Ziena Hassan Razaq Al Hadad a, "Embedding Data in Non-Important Gabor Ridges," 2022.
[3] A. P. D. A. A. M. Q. a. M. s. A. Y. Abdulkadhim2, "Comparison study about Car Detection and Type Recognition Techniques," 2022.
[4] A. A. M. B. Qazzaz and N. E. Kadhim, "Watermark Based on Singular Value Decomposition," Baghdad Science Journal, 2023.
[5] A Anand and AK Singh. A comprehensive study of deep learning based covert communication. ACM Transactions on Multimedia Computing, Communications, and Applications, 18(2s):1– 19, 2022.
[6] AK Singh. 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, 2019.
[7] SP Mohanty et al. Everything you want to know about watermarking: from paper marks to hardware protection. IEEE Consumer Electronics Magazine, 6(3):83–91, 2017.
[8] S Gull et al. An efficient watermarking technique for tamper detection and localization of medical images. Journal of Ambient Intelligence and Humanized Computing, 11(5):1799–1808, 2020
[9] D Singh and SK Singh. Dct based efficient fragile watermarking scheme for image authenti- cation and restoration. Multimedia Tools and Applications, 76(1):953–977, 2017
[10] XT Wang et al. Reversible data hiding for high quality images exploiting interpolation and direction order mechanism. Digital Signal Processing, 23(2):569–577, 2013.
[11] NS Kamaruddin et al. A review of text watermarking: theory, methods, and applications. IEEE Access, 6:8011–8028, 2018.
[12] P Amrit and AK Singh. Survey on watermarking methods in the artificial intelligence domain and beyond.

Computer Communications, 188:52–65, 2022.

[13] Y Li, H Wang, and M Barni. A survey of deep neural network watermarking techniques. Neurocomputing, 461:171–193, 2021.

[14] W Wan et al. A comprehensive survey on robust image watermarking. Neurocomputing, 488:226– 247, 2022.

[15] Y Li, H Wang, and M Barni. A survey of deep neural network watermarking techniques. arXiv preprint arXiv:2103.09274, 2021.

[16] O Byrnes et al. Data hiding with deep learning: a survey unifying digital watermarking and steganography. arXiv preprint arXiv:2107.09287, 2021.

[17] watermark detection and removal with deep convolutional networks. In Pat- tern Recognition and Computer Vision: First Chinese Conference, PRCV 2018, Guangzhou, China, November 23-26, 2018, Proceedings, Part III 1, pages 27–40. Springer, 2018.

[18] M Bagheri et al. Adaptive control of embedding strength in image watermarking using neural networks. arXiv preprint arXiv:2001.03251, 2020.

[19] W. Zheng et al. Robust and high capacity watermarking for image based on dwt-svdand cnn. In 13th IEEE Conf. Ind. Electron. and Appl., pages 1233–1237. IEEE, May 2018.

[20] D. Li et al. A novel cnn based security guaranteed image watermarking generation scenario for smart city applications. Information Sciences, 479:432–447, 2019.

[21] M. Plata and P. Syga. Robust spatial-spread deep neural image watermarking. In IEEE 19th Int. Conf. Trust, Security and Privacy in Comput. and Commun. (TrustCom), pages 62–70. IEEE,December 2020.

[22] S. Mun et al. A robust blind watermarking using convolutional neural network. arXiv preprint arXiv:1704.03248, 2017.

[23] A. K. Singh, M. Dave, and A. Mohan. Hybrid technique for robust and imperceptible multiple watermarking using medical images. Multimedia Tools and Applications, 75(14):8381–8401, 2016.

[24] S. Thakur et al. Improved dwt-svd-based medical image watermarking through hamming code and chaotic encryption. In D. Dutta et al., editors, Advances in VLSI, Communication, and Signal Processing, pages 897–905. Springer, 2020.

[25] A. Anand and A. K. Singh. An improved dwt-svd domain watermarking for medical informa- tion security. Computer Communications, 152:72–80, 2020.

[26] A. Anand et al. Compression-then-encryption-based secure watermarking technique for smart healthcare system. IEEE Multimedia, 27(4):133–143, 2020.

[27] S Ingaleshwar and NV Dharwadkar. Water chaotic fruit fly optimization-based deep convolu- tional neural network for image watermarking using wavelet transform. Multimedia Tools and Applications, pages 1–25, 2021.

[28] O. P. Singh and A. K. Singh. Data hiding in encryption–compression domain. Complex Intelligent Systems, pages 1–14, 2021.

[29] L. Y. Hsu and H. T. Hu. Qdct-based blind color image watermarking with aid of gwo and dncnn for performance improvement. IEEE Access, 9:155138–155152, 2021.

[30] H. Kandi, D. Mishra, and S. R. S. Gorthi. Exploring the learning capabilities of convolutional neural networks for robust image watermarking. Computers & Security, 65:247–268, 2017.

[31] Y. Nagai et al. Digital watermarking for deep neural networks. International Journal of Multimedia Information Retrieval, 7(1):3–16, 2018.

[32]X. Guan et al. Reversible watermarking in deep convolutional neural networks for integrity authentication. In Proc. 28th ACM Int. Conf. Multimedia, pages 2273–2280, 2020.

[33] Q. Wei, H. Wang, and G. Zhang. A robust image watermarking approach using cycle varia- tional autoencoder. Security Commun. Netw., 9:2020, 2020.

[34] L. Zhang, W. Li, and H. Ye. A blind watermarking system based on deep learning model. In IEEE 20th Int. Conf. Trust, Security, and Privacy in Comput. and Commun. (TrustCom), pages 1208–1213, 2021.

[35] B. Fan, Z. Li, and J. Gao. Dwimark: a multiscale robust deep watermarking framework for diffusion-weighted imaging images. Multimedia Syst., 28(1):295–310, 2022.

[36] H. Fang et al. Encoded feature enhancement in watermarking network for distortion in real scenes. IEEE Trans. Multimedia, pages 1–13, 2022.

[37] X. Kang, J. Huang, and W. Zeng. Efficient general print-scanning resilient data hiding based on uniform log-polar mapping. IEEE Trans. Inf. Forensics Security, 5(1):1–12, 2010.
[38] J. Zhu et al. Hidden: hiding data with deep networks. In Lect. Notes Comput. Sci., volume 11219, pages 657–672, 2018.
[39] Y. Liu et al. A novel two-stage separable deep learning framework for practical blind water-marking. In Proc.27th ACM Int. Conf. Multimedia, pages 1509–1517, 2019.
[40] Z. Ma et al. Local geometric distortions resilient watermarking scheme based on symmetry. IEEE Trans. Circuits Syst. Video Technol., 31(12):4826–4839, 2021.
[41] P. Neekhara et al. Facesigns: semi-fragile neural watermarks for media authentication and countering deepfakes. arXiv:2204.01960, 2022.
[42] J. Wu et al. De-mark gan: removing dense watermark with generative adversarial network. In Int. Conf. Biometrics, IEEE, pages 69–74, 2018.
[43] V. L. Cu et al. A robust data hiding scheme using generated content for securing genuine documents. In Int. Conf. Document Anal. and Recognit., IEEE, pages 787–792, 2019.
[44] C. V. Loc, J. C. Burie, and J. M. Ogier. Stable regions and object fill-based approach for document images watermarking. In 13th IAPR Int. Workshop on Document Anal. Syst., April, IEEE, pages 181–186, 2018.
[45] C. V. Loc, J. C. Burie, and J. M. Ogier. Document images watermarking for security issue using fully convolutional networks. In 24th Int. Conf. Pattern Recognit., August, IEEE, pages 1091–1096, 2018.
[46] V. L. Cu, J. C. Burie, and J. M. Ogier. Watermarking for security issue of handwritten docu- ments with fully convolutional networks. In 16th Int. Conf. Front. in Handwriting Recognit., August, IEEE, pages 303–308, 2018.
[47]J. Hou et al. Reversible data hiding based on multiple histograms modification and deep neural networks. Signal Process. Image Commun., 92:116–118, 2021.
[48] W. He et al. Efficient pvo-based reversible data hiding using multistage blocking and prediction accuracy matrix. J. Visual Commun. Image Represent., 46:58–69, 2017.
[49]Y. Jia et al. Reversible data hiding based on reducing invalid shifting of pixels in histogram shifting. Signal Process., 163:238–246, 2019.
[50] X. Li et al. Efficient reversible data hiding based on multiple histograms modification. IEEE Trans. Inf. Forensics Security, 10(9):1824–1834, 2015.
[51] J. Zhang et al. Protecting intellectual property of deep neural networks with watermarking. In Proc. 2018 on Asia Conf. Comput. and Commun. Security, pages 159–172, 2018.
[52] H. Wu et al. Watermarking neural networks with watermarked images. IEEE Trans. Circuits Syst. Video Technol., 31(7):2591–2601, 2021.
[53] F. Deeba et al. Digital watermarking using deep neural networks. Int. J. Mach. Learn. Comput., 10(2):277–282, 2020.
[54]I. Hamamoto and M. Kawamura. Image watermarking technique using embedder and extractor neural networks. IEICE Trans. Inf. Syst., E102.D(1):19–30, 2019.
[55]Xiang Li, Chan Lu, Danni Cheng, Wei-Hong Li, Mei Cao, Bo Liu, Jiechao Ma, and Wei- Shi Zheng. Towards photorealistic visible watermark removal with conditional generative adversarial networks. In ICIG, 2019.
[56]Yang Liu, Zhen Zhu, and Xiang Bai. Wdnet: Watermark decomposition network for visible watermark removal. In WACV, 2021.
[57] Xiaodong Cun and Chi-Man Pun. Split then refine: stacked attention-guided resunets for blind single image visible watermark removal. In AAAI, 2021.
[58] Jing Liang, Li Niu, Fengjun Guo, Teng Long, and Liqing Zhang. Visible watermark removal via self-calibrated localization and background refinement. In ACMMM, 2021.
[59] Lijun Fu, Bei Shi, Ling Sun, Jiawen Zeng, Deyun Chen, Hongwei Zhao, and Chunwei Tian. An improved u-net for watermark removal. Electronics, 11(22):3760, 2022.
[60] Xinwei Liu, Jian Liu, Yang Bai, Jindong Gu, Tao Chen, Xiaojun Jia, and Xiaochun Cao. Watermark vaccine: Adversarial attacks to prevent watermark removal. In ECCV, 2022.

[61] Cong Wang, Yutong Wu, Zhixun Su, and Junyang Chen. Joint self-attention and scale- aggregation for self-calibrated deraining network. In ACMMM, 2020.
[62] Yves Grandvalet and Yoshua Bengio. Semi-supervised learning by entropy minimization. In NIPS, 2004.
[63] Xiaodong Cun, Chi-Man Pun, and Cheng Shi. Towards ghost-free shadow removal via dual hierarchical aggregation network and shadow matting gan. In AAAI, 2020.
[64] Bin Ding, Chengjiang Long, Ling Zhang, and Chunxia Xiao. Argan: Attentive recurrent generative adversarial network for shadow detection and removal. In ICCV, 2019.
[65] Zhendong Wang, Xiaodong Cun, Jianmin Bao, Wengang Zhou, Jianzhuang Liu, and Houqiang Li. Uformer: A general u-shaped transformer for image restoration. In CVPR, 2022.
[66] Junlin Han, Weihao Li, Pengfei Fang, Chunyi Sun, Jie Hong, Mohammad Ali Armin, Lars Petersson, and Hongdong Li. Blind image decomposition. In ECCV, 2022.
[67] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In CVPR, 2021.
[68] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-end object detection with transformers. In ECCV, 2020.
[69] Yuqing Wang, Zhaoliang Xu, Xinlong Wang, Chunhua Shen, Baoshan Cheng, Hao Shen, and Huaxia Xia. End-to-end video instance segmentation with transformers. In *CVPR*, 2021.
[70]Lin Huang, Jianchao Tan, Ji Liu, and Junsong Yuan. Hand transformer: non-autoregressive structured modeling for 3d hand pose estimation. In ECCV, 2020.
Fuzhi Yang, Huan Yang, Jianlong Fu, Hongtao Lu, and Baining Guo. Learning texture transformer network for image super-resolution. In CVPR, 2020.
[71] Hanting Chen, Yunhe Wang, Tianyu Guo, Chang Xu, Yiping Deng, Zhenhua Liu, Siwei Ma, Chunjing Xu, Chao Xu, and Wen Gao. Pre-trained image processing transformer. In CVPR, 2021.
[72] Xiaofeng Cong, Jie Gui, Kai-Chao Miao, Jun Zhang, Bing Wang, and Peng Chen. Discrete haze level dehazing network. In ACMMM, 2020.
[73] Adam W Harley, Konstantinos G Derpanis, and Iasonas Kokkinos. Segmentation-aware con- volutional networks using local attention masks. In ICCV, 2017.
[74] Martin Simonovsky and Nikos Komodakis. Dynamic edge-conditioned filters in convolutional neural networks on graphs. In CVPR, 2017.
[75] HyunJae Lee, Hyo-Eun Kim, and Hyeonseob Nam. Srm: A style-based recalibration module for convolutional neural networks. In ICCV, 2019.
[76] Z. Cao, S. Niu, J. Zhang, and X. Wang. Generative adversarial networks model for visible watermark removal. IET Image Processing, 13(10):1783–1789, 2019.
[77] Yizeng Han, Gao Huang, Shiji Song, Le Yang, Honghui Wang, and Yulin Wang. Dynamic neural networks: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.
[78] Xing Zhao, Li Niu, and Liqing Zhang. Visible watermark removal with dynamic kernel and semantic-aware propagation. In BMVC, 2022.