Grenze International Journal of Engineering and Technology, Jan Issue

Watermarking Techniques Based on Deep Learning Approaches: A Review

Bharathi Chidirala[1] and Bibhudendra Acharya[2]
[1-2]Department of ECE, National Institute of Technology Raipur, India
Email: {bchidirala.phd2022.etc, bacharya.etc}@nitrr.ac.in

***Abstract*— Digital watermarking plays a vital role in protecting the rights and integrity of images, preventing unauthorized use, and promoting a safer and more accountable digital environment. Conventional watermarking schemes often suffer from limited robustness and susceptibility to various image processing operations, while deep learning-based watermarking schemes provide enhanced resilience to attacks and greater adaptability to complex image transformations, making them more effective in maintaining the integrity and ownership of digital content. This paper provides a concise analysis of commonly utilized deep learning networks and watermarking methods. The article begins by discussing the general framework, applications, and classification of watermarking based on different use cases. It then delves into deep learning network models such as Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN), and Deep Neural Networks (DNN), along with associated watermarking techniques. Lastly, the paper addresses the challenges and offers guidance for future developers and researchers in this field.**

Index Terms— **Digital watermarking, deep learning, data hiding, encoder, decoder, CNN, GAN, DNN**

I. INTRODUCTION

Digital watermarking is a popular technique that involves the embedding of hidden data, referred to as a watermark, into a digital image. This watermark can be thought of as a unique digital signature that is imperceptibly embedded within the host or cover image. This process is typically performed using specialized algorithms [1], [2]. The primary purpose of digital watermarking is to ensure the confidentiality and integrity of transmitted information. By embedding a watermark into a digital image, it becomes possible to establish ownership and protect the copyright of the content. Watermarks serve as a form of digital identification, enabling content creators and owners to prove their rights over the material. It also facilitates authentication using the watermarking algorithms particularly useful in scenarios where tampering or unauthorized modifications need to be detected.

In the process of watermarking, there are three primary goals that need improvement: imperceptibility, capacity, and robustness. Robustness evaluates the watermark's ability to withstand different attacks by comparing extracted watermark to the original one. Imperceptibility, on the other hand, refers to how similar the watermarked image appears compared to that of original/host image. Striking a balance between invisibility and robustness is crucial since enhancing invisibility can weaken robustness. The capacity of a watermark refers to the no. of bits that can be inserted into the host/cover image. The information capacity of a carrier image depends on its ability to hold data, and expanding the capacity diminishes the level of imperceptibility of the watermark. Figure 1 illustrates the

Grenze ID: 01.GIJET.10.1.69

watermarking trade-off that outlines the required specifications for any watermarking approach. It is essential for watermarking techniques to be compatible with various media and find applications in diverse fields.
The fundamental process of digital watermarking is illustrated in Figure 2. During the embedding stage, a user-specific algorithm is employed to implant the watermark onto the original or host image. The watermarked image is then transmitted through a noisy channel. At the receiving end, the watermarked image is received with noise. Using the same method and keys employed during transmission, the watermarks are extracted. This enables the secure transmission of private data between the sender and receiver over unprotected channels and connected devices, without significantly distorting the host data.
Here are the major applications of digital watermarking

- *Copyright Protection*: Embedding watermarks to deter unauthorized copying and piracy of digital content.
- *Authentication and Tamper Detection*: Verifying the authenticity, integrity of digital files, and preventing counterfeiting of important documents and images, detecting any modifications or tampering attempts.
- *Broadcast Monitoring*: Identifying and tracking copyrighted material during television or radio broadcasts.
- Content Identification and Metadata: Embedding metadata and identification information for efficient content management and retrieval.
- *Transaction Tracking and Fingerprints*: Using watermarks as unique identifiers to track digital content distribution and monitor usage.
- *Multimedia Forensics*: Analyzing and identifying manipulated or altered digital content for forensic investigations.

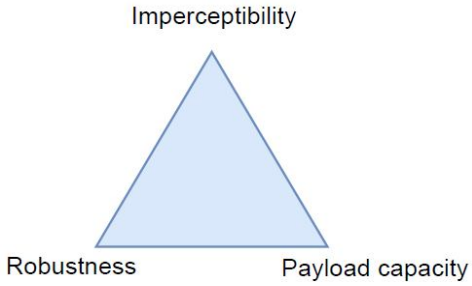

Figure 1. Tradeoff among imperceptibility, payload capacity and robustness

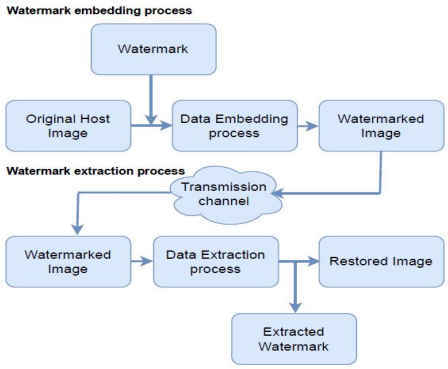

Figure 2. General framework of data embedding and extraction

These applications highlight the versatility and importance of digital watermarking in protecting intellectual property rights, ensuring content integrity, and enabling secure digital transactions.
The following are the contributions of our article:

- Firstly discussed about the basic concept of digital watermarking and trade-off between level of imperceptibility, payload capacity and robustness, recent applications, and digital watermarking methodology.
- Then, classification of digital watermarking and significance of deep learning over traditional approaches is been discussed.
- Basic architectures of CNN, DNN, GAN networks are presented with suitable diagrams and explained how these networks are designed for image processing tasks. We review the recent deep learning based watermarking schemes.

- Few issues related to deep learning models are discussed and widely used performance evaluation metrics are also mentioned.

The paper is organized as follows: Firstly digital watermarking concept, trade off among imperceptibility, robustness and payload capacity, general frame work of data embedding and extracting, few applications of watermarking are discussed. In Section II classification based on different types of applications are discussed. In Section III various models in deep learning and summary on deep learning based watermarking schemes is being discussed. Section IV describes few issues that are identified. In Section V mentioned few performance metrics by which we can find out the efficiency of the scheme and finally in Section VI the conclusions drawn are presented.

II. Classification Of Digital Watermarking

Based on its applications as well as requirement, watermarking techniques may be divided into many categories. This can be illustrated in Figure 3.

A. Based on the host signal

The watermarking base can be classified into text, picture, audio, as well as video categories depending on the cover image or host signal. The MPEG audio encoding, filtering, resampling, and cropping attacks are only a few examples of signal processing attacks that the watermarking approach should be resistant to.

B. Based on the perceptibility

The watermarking content which is visible to our naked eye is the visible watermarking and which are not visible (hidden) in the original content is called invisible watermarking. Applications that employ invisible watermarking to offer authentication when delivering the secret message also weuse visible watermarking techniques, such as those used on bank notes and business logos, to give ownership and copy right protection.

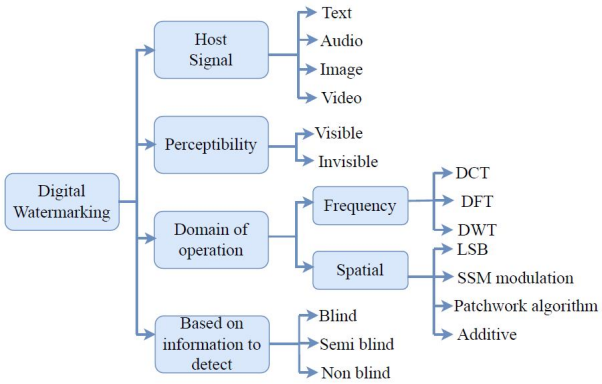

Figure 3. Classification of digital watermarking [5]

C. Based on the information to detect

i) Blind watermarking uses no original signal to find the embedded data in the watermarked image. These blind watermarking methods are weaker against most of the attacks [3]. These also go by the name of public watermarking method. ii) Semi-blind watermarking, which does not use the original signal and requires some keying expertise to discover the encoded data from a watermarked image. These methods of watermarking are often known as semi-private approaches. iii) Non-blind watermarking identifies the embedded watermark using original signal and keys as well as any other information that is accessible at the transmission end. These methods are often known as private watermarking methods. These methods are more resistant to numerous threats than the other two methods.

D. Based on the domain of operation

i) Frequency domain watermarking means the pixels are transformed in to their spectral coefficients and then data is embedded. ii) In Spatial domain watermarking technique, the alterations done directly on the pixels. The effectiveness of these approaches depends on their robustness and imperceptibility. The simplest use of a spatial domain approach is the alteration of the least significant bit, whereas some applications of a frequency domain technique include the adjustment of DCT and DWT coefficients.

E. Significance of deep learning over traditional approaches

The adoption of deep learning approaches specifically CNN based techniques rather than transformed-based approaches for watermarking is very advantageous because of their powerful capacity for learning and improved, more accurate outputs [9].

- *Robustness*: Deep learning algorithms can embed watermarks in a way that is robust to common image processing operations, like cropping, resizing, and compression, making it difficult to clear the watermark without affecting the quality of the image.
- *Detection*: Deep learning algorithms can also be used for watermark detection, allowing for more accurate and efficient detection of watermarks compared to traditional methods.
- *Complexity*: Deep learning algorithms can handle complex watermarking scenarios that may not be possible with traditional methods, such as embedding multiple watermarks in a single image or embedding watermarks in different frequency domains.
- *Generalization*: Deep learning algorithms can generalize well to new images, making them well-suited for watermarking in large-scale applications.

Overall, deep learning has the potential to significantly improve the effectiveness, robustness, and efficiency of digital watermarking algorithms compared to traditional methods.

III. Overview Of Recent Deep Learning Based Watermarking Schemes

In recent years, several surveys on digital media protection, hardware, and artificial model protection employing watermarking approaches have been presented. In [4] the functions and use of deep-learning models in the various stages of the watermarking approaches were summarized by the authors. In [5] authors addressed about several artificial intelligence watermarking techniques, beyond those few machine learning based and deep learning based watermarking techniques were also discussed. A thorough investigation of watermarking's use to secure deep learning models is summarized. In [6] delivered deep learning surveys on intellectual property protection. The main intention of the study, in contrast, is to give a precise summary of the salient features of deep-learning models, which are frequently utilized in watermarking for ownership, model protection, and copyright protection. This article communicates few recently published proposals in deep learning model.

Without the need for feature representations or extractions, deep learning-based frameworks automatically learn from training data to represent the hierarchical data [7]. Deep learning models don't need to manually represent features in their learning models. To be more precise, a deep network maps the material that needs to be processed and accepts it as input in its raw format (such as an image or an audio signal). They have recently been widely employed in data concealing and image processing due to their amazing capacity to replicate the learning capabilities of the human brain and interact more naturally [8]. Deep-learning models like the CNN, DNN, GAN, Recurrent Neural Network, and Artificial Neural Networks are often utilized in watermarking methods as shown in Figure 4.

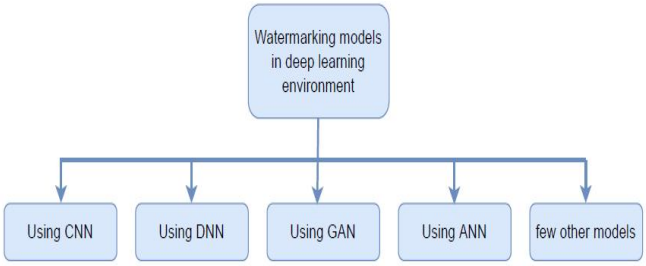

Figure 4. Watermarking models in deep learning environment [16]

A. Convolutional Neural Networks (CNNs), Deep Neural Networks (DNNs), and Generative Adversarial Networks (GANs)

CNN, DNN, GAN are main types of neural networks used in the field of machine learning and deep learning. However, they have distinct characteristics and are used for different purposes.

1) CNNs are a type of neural networks specifically designed for image processing tasks. Figure 5 shows the basic architecture of this network. It basically consists of 5 layers. (i) Input layer, (ii) convolution layer, (iii) pooling layer, (iv) fully connected layer and (v) output layer. They use convolutional layers to extract the features from given images and classify them into different categories. CNNs are commonly used for tasks such as object detection, image classification, and semantic segmentation. To determine its role in watermarking, these can be

used for robust watermarking, by embedding the watermark into image in such a way that it is difficult to remove without significantly affecting the quality of the image. Additionally, CNNs can also be used for watermark detection, by training the network to recognize the presence of the watermark in an image and to differentiate it from other similar patterns in the image [10], [11]. Overall, deep learning has the potential to significantly improve the effectiveness and robustness of digital watermarking algorithms.

2) DNNs are neural networks with multiple hidden layers as shown in figure 6 that can learn complex non-linear relationships between inputs and outputs. The more the hidden layers the higher the time complexity. DNNs can be used for a different variety of tasks such as Image classification, Natural language processing (NLP) and speech recognition. The DNN can be trained in a supervised manner, using a dataset of watermarked and original images, to optimize the watermarking and detection process [12]. Once trained, the DNN can be used to watermark new images and detect the presence of watermarks in a fast and efficient manner. Watermarking based on DNNs has the potential to significantly improve the robustness, effectiveness, and efficiency of digital watermarking algorithms compared to traditional methods [13]. However, it requires computational resources and huge amount of training data. This may limit its practical application in certain scenarios.

3) GANs are a type of neural network used for generative tasks, such as image synthesis and style transfer. GAN consists of two neural networks, one is generator and other is discriminator, that are trained in a competitive manner. The architecture of GAN model is shown in figure 7. The generator produces fresh images, while the discriminator achieves to differentiate between the generated images and real images. [14]. These networks can be used to help in watermarking by using the adversarial loss function to train a generator network to insert watermarks into images, while a discriminator network is trained to detect the presence of the watermarks. During training, the generator network learns to produce images that are identical from the original images with the added watermark, while the discriminator network learns to detect the presence of the watermark [14], [15]. This results in a generator network that is capable of adding watermarks to images in such a way that they are difficult to remove without significantly altering the image.

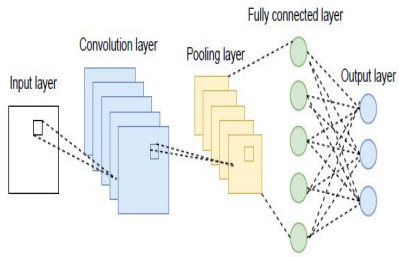

Figure 5. Basic architecture of CNN

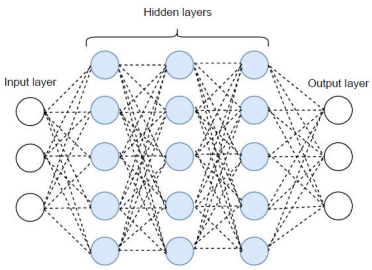

Figure 6. Basic architecture of CNN

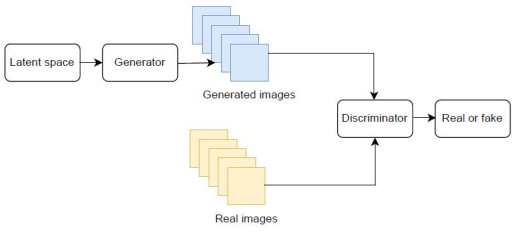

Figure 7. Basic architecture of GAN

The steps involved in data hiding using deep learning model are: i) The identification of embedding location in cover content, ii) selection of appropriate embedding strength, by which it efficiently offers a equalized trade-off between accuracy, quality and robustness, iii) Simulation of attacks for efficient watermark extraction, iv) Error reduction and mark extraction denoising, and v) Copyright protection and deep learning model ownership verification. HK Singh and AK Singh in [16] gave a detailed description and classifications of deep learning based watermarking techniques.

In order to enhance the tradeoffs of the watermark, Kandi et al. [17] suggested a non-blind categorized transform domain based watermarking approach employing the CNN network. To increase the techniques resilience, a CNNs auto-encoder learning capacity is used. It embeds utilizing the CNN weights, which are quite vulnerable to assaults, using the input-output data.

A GAN based deep-learning technique was suggested by Wu et al. [18] to validate the distorted images brought on by thick watermarks. For enhancing the quality and verification performance of the recovered pictures, the model comprises of a generator and a discriminator. The generator is an autoencoder that converts damaged, heavily watermarked pictures into a representation vector and then decodes the vector into an RGB image. The discriminator regulates how the resulting pictures are composed and reduces feature loss. The recovered pictures features and ground truth images are extracted using the ResNet-46 model. At a false positive rate (FPR) of 1%, it obtains a verification accuracy of 96.36%.

Bagheri et al. [19] proposed a blind image watermarking approach in which the embedding strength was computed using Mask RCNN. By computing the block into the chosen discrete wavelet transform as well as discrete cosine transform blocks using lower region-of-interest pixels, the watermark was inserted in the cover image. Though the watermark is successfully recovered, the approach only offers a little amount of embedding capacity and watermark security.

To test the watermarks imperceptibility and resilience, Zheng et al. [20] created a wavelet based watermarking technique for digital material. The watermark data was first added into high band of cover image after it had been first converted into various bands using DWT. The watermark sequence was then included into the chosen low bands after the wavelet transformation was done to the low bands. The watermarks singular vector was used to create the watermark sequence, and the scrambled sequence was used to create the singular. Later, a strong relationship between the watermark, cover media, and watermarked material was formed using CNN to enable watermark extraction. They are more suitable to many applications because of their affordable embedding costs for various sorts of watermarks.

In 2021, Mellimi et al. [21] proposed an effective image watermarking technique based on the DNN and LWT (lifting wavelet transform). Wavelet transforms are used in this scheme to embed watermark. It also maintained good level of imperceptibility and robustness. In order to discover the best tradeoff between robustness and achievable imperceptibility, various frequency bands are examined. Various techniques, including noise attacks, compression attacks, and filtering attacks, are utilised to test the robustness of the suggested work. The DNN is then trained to recognise the modifications that these attacks make to the various bands of frequency. The invisibility of the system is improved by the employment of the high-frequency sub-bands like LH / LH1 / HL for watermark insertion. The proposed technique was tested on 600 images, and the suggested method's PSNR was 44.11 in dB, with the variance of 0.51 and a standard deviation of 0.71. Good performance is observed when compared with existing methods [22], [23].

Ge et al. [24] presented a model based on DNN for end-to-end document image watermarking system. The watermark is embedded and extracted using encoders and decoders, respectively. To simulate the numerous attacks that can actually occur, such as the Dropout, Cropout, Blurring, Gaussian noise, Image resize, and JPEG compression, a noise layer is introduced. To keep character embedding alteration to a minimum, a text-sensitive loss function was created. An approach for adjusting the embedding strength is suggested to enhance the watermarked image's quality while minimizing the impact on extraction precision. As there was the shortage in ready made data sets for document images, the authors collected two large scale

datasets. One with English sentences named DocImgEN and the other with Chinese sentences DocImgCN. Each set contained a total of 250,000 document images. Out of these, 230,000 images were utilized for training, 10,000 for validation, and the remaining 10,000 images for testing. The experimental results were then compared to three state-of-the-art schemes [25]–[27] which has shown better performance.

Mahapatra et al. [28] in 2022 presented a CNN based watermarking scheme using auto encoder model. Firstly extracted features of host/cover image and watermark image and concatenated. In the extraction Relu algorithm and transposed block level convolution is applied to extract the features. It is tested for different attacks like salt and pepper noise, speckle, rotation, gaussian, and JPEG compression. The authors compared their results with few other schemes [29], [30] and shown better robustness.

Using DNN [31] proposed a transparent watermarking scheme for videos. Authors presented a technique for watermark embedding using neural networks when using the chrominance channel of the YUV420p colour model for lossy video compression with the HEVC codec. The training procedure for DNN to include a watermark while the compression channel is there. This method is distinguished by the embedded watermark's high accuracy when compared to the original video. Within the range/spectrum of compression values specified by the CRF (which takes context into account), the approach allows for the entire watermark extraction from the single video frame that has been compressed using this HEVC codec.

Fei et al. [32] presented a model to watermark the GAN model directly so that the images generated by this model gets watermarked to protect the copyright information. By using fine-tuning to embedd the watermark, the computational complexity is significantly reduced, enabling GAN owners to quickly safeguard already trained GAN models. The suggested approach can also serve as a positive approach to GAN image identification and attribution; By analyzing the existence of watermarks embedded within the image, it becomes possible to associate the image with the specific GAN model responsible for its creation. Average bit accuracy for PGGAN and styleGAN2 models with and without augmentation are 99.33 and 99.44 respectively. Compared their results with other existing state-of-art scheme [33] and got higher bit accuracy. Table I shows summary of discussed watermarking schemes based on deep learning.

TABLE I. SUMMARY ON DEEP LEARNING BASED WATERMARKING SCHEMES

Ref No.	Objective	Goal	Approaches	Model used	Cover image/ Watermark size	Results	Applications	Limitations observed
[17]	To increase imperceptibility and robustness	Authenticated against number of image attacks	Autoencoder	CNN	128 x128 / 64 x 64	PSNR when no noise is induced is 58.91, BER = 0 and SSIM = 0.9997	Grayscale general images	Can be extended to other watermarking categories and host signals
[21]	Effective image watermarking and fast processing during extraction and high robustness	Fast embedding and extracting is achieved	LWT	DNN	512 x 512 / 32 x 32	PSNR = 44.1148 dB, variance = 0.5146 with the standard deviation of 0.7173	Gray scale Images	Performance of the attacks like SN (1%) and SPN (1%) are to be improved
[24]	To design textsensitive loss function to decrease the modification on text characters and to improve the robustness	Watermarking in text / document images with increased algorithm strength	Encoder and Decoder	DNN	400 x 400	PSNR, SSIM, CPP are 40.10, 0.962, 3.52 for DocImgEN and PSNR, SSIM, CPP are 41.07, 0.965, 3.77 for DocImgCN	Document Images	Little more efforts are required to resist JPEG compression Attack
[28]	To maximize the robustness and also to assure invisibility	Robustness and imperceptibility improvement	Autoencoder	CNN	128 x 128 / 64 x 64	PSNR = 31.34 dB, SSIM = 0.9940 and NC = 0.9937	Grayscale general images	Scheme could have concentrated a little more on improving the visual quality (PSNR)
[31]	High efficiency video watermarki	This process of watermarking videos	Adjustable subsquares properties algorithm	DNN and ANN	frame size 128 x 128 / 96 bits	PSNR = 47.299 dB, time for embedding	Video Watermarking	More investigation can be done to

	ng is their main objective while ensuring the video quality	ensures that the quality of the watermarks is maintained.				and extracting for a single frame which has 416 x 240 resolution was 81.132 & 82.223 sec		determine whether it is possible to reproduce an image with embedded watermarks that is more faithful.
[32]	To watermark the GAN model so that any image generated by this model gets watermarked to protect the copyright information	Watermarking to secure complete network	Encoder and Decoder	GAN	256 x 256 / 100 bits and 128 x 128 / 50 bits	Avg. bit accuracy for PGGAN and style-GAN2 models with and without augmentation are 99.33 and 99.44 respectively	GAN network	Watermarking capacity is observed to be less. One can research on increasing its capacity without much variation in visual quality

IV. Few Issues Related To Deep Learning Models

1. The majority of approaches did not significantly address complexity problems, deep learning model security, or data security.
2. Studies have indicated that employing transform domain methods for watermarking data is more dependable compared to spatial domain methods. Hence, it becomes crucial to overcome the constraints of relying solely on a single domain approach when implementing deep learning-based watermarking techniques.
3. Pre-trained models are frequently employed to reduce the complexity of model training, however this practice creates the issues of model overwriting and surrogate model attacks.
4. The flexibility of the watermarking system is greatly influenced by the number of samples utilized for training and the choice of loss function.
5. The embedding capability of the majority of deep learning-based watermarking is low.
6. Trade-off between embedding capacity and imperceptibility remains challenging.

V. Performance Metrics For Digital Watermarking

To avoid distortion, maintain robustness and imperceptibility original image and embedded image are to be similar and should be seen without any difference. Hereby few metrics and equations are mentioned by which we can calculate the embedding scheme's performance.

A. PSNR: The Peak Signal-to-Noise Ratio (PSNR) represents the ratio between the maximum potential power of an image and the maximum power of corrupting noise that affects its quality. It can be quantified using (1)

$$PSNR = 10 \times log_{10}(\frac{255^2}{MSE}) \quad (1)$$

B. MSE: It means the mean of squared difference between the expected parameter and the actual parameter. This is calculated by (2)

$$MSE = \frac{1}{mn}\sum_{x=0}^{m-1}\sum_{y=0}^{n-1}[I(x,y) - I_w(x,y)]^2 \quad (2)$$

In the equation, "m" represents the number of rows, "n" denotes the number of columns, and (x, y) represents the pixel index within the image.

C. SSIM: The SSIM is the measure of visual metric which identifies the image quality degradation when data is compressed or made any changes into the original image and can be calculated by the help of below eqn.

$$SSIM(I, I_w) = \frac{(2\mu_I \mu_{I_w} + C_1) + (2\sigma_{II_w} + C_2)}{(\mu_I^2 + \mu_{I_w}^2 + C_1)(\sigma_I^2 + \sigma_{I_w}^2 + C_2)} \quad (3)$$

Where I, Iw and I, Iw are mean and variance values of original and embedded images. C1 and C2 are the predefined offset values.

D. BER: It can be defined as the ratio of falsely decoded bits to the total number of transmitted bits.

$$BER = \frac{numberofinaccuratelydecodedbits}{totalnumberoftransmittedbits} \quad (4)$$

Figure 8 shows the sample images for simulation. In most of the literature survey these are the commonly used images in image processing for experimentation. The PSNR values of the scheme [34] ranged from 47.63 dB to 50.62 dB. It shows that the embedded watermark cannot be detected. And the PSNR with just noticeable difference ranges between 28.51 dB and 31.43 dB.

Figure 8. Test images [34]

VI. CONCLUSION

Significant advancements have been achieved in the field of image processing through deep learning. More recently, deep learning has also shown progress in the domain of data concealment, offering an efficient and reliable watermarking approach. Based on the well-known deep-learning model utilized for the watermarking techniques, this work offers a brief assessment. Beginning with watermarking applications and classification, we listed few mostly used deep learning networks and few watermarking schemes. The methods, model function, results, comparative papers of the deep-learning-based were all discussed. We also discussed a few drawbacks of deep learning-based methods. We believe that this survey will provide useful information for aspiring researchers and provide insight into applying the deep-learning model for watermarking approaches.

REFERENCES

[1] Cox, Ingemar, et al. “Digital watermarking and steganography.” Morgan kaufmann, 2007.

[2] Lu, Chun-Shien, ed. “Multimedia security: steganography and digital watermarking techniques for protection of intellectual property.” Igi Global, 2005.

[3] Uccheddu, Francesca, Massimiliano Corsini, and Mauro Barni. “Wavelet-based blind watermarking of 3D models.” Proceedings of the 2004 workshop on Multimedia and security. 2004.

[4] Anand, Ashima, and Amit Kumar Singh. “A Comprehensive Study of Deep Learning-based Covert Communication.” ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 18.2s (2022): 1-19.

[5] Amrit, Preetam, and Amit Kumar Singh. “Survey on watermarking methods in the artificial intelligence domain and beyond.” Computer Communications 188 (2022): 52-65.

[6] Byrnes, Olivia, et al. “Data hiding with deep learning: A survey unifying digital watermarking and steganography.” arXiv preprint arXiv:2107.09287 (2021).
[7] Heaton, Jeff. “Ian goodfellow, yoshua bengio, and aaron courville: Deep learning.” (2018): 305-307.
[8] Hatoum, Makram W., et al. “Using deep learning for image watermarking attack.” Signal Processing: Image Communication 90 (2021): 116019.
[9] Vukoti, Vedran, Vivien Chappelier, and Teddy Furon. “Are deep neural networks good for blind image watermarking?.” 2018 IEEE International Workshop on Information Forensics and Security (WIFS). IEEE, 2018.
[10] Li, Daming, et al. “A novel CNN based security guaranteed image watermarking generation scenario for smart city applications.” Information Sciences 479 (2019): 432-447.
[11] Dhaya, R. “Light weight CNN based robust image watermarking scheme for security.” Journal of information technology and digital world 3.2 (2021): 118-132.
[12] Li, Zheng, et al. “How to prove your model belongs to you: A blind-watermark based framework to protect intellectual property of DNN.” Proceedings of the 35th Annual Computer Security Applications Conference. 2019.
[13] Jia, Zhaoyang, Han Fang, and Weiming Zhang. “Mbrs: Enhancing robustness of dnn-based watermarking by mini-batch of real and simulated jpeg compression.” Proceedings of the 29th ACM international conference on multimedia. 2021.
[14] Hao, Kangli, Guorui Feng, and Xinpeng Zhang. “Robust image watermarking based on generative adversarial network.” China Communications 17.11 (2020): 131-140.
[15] Li, Qi, et al. “Concealed attack for robust watermarking based on generative model and perceptual loss.” IEEE Transactions on Circuits and Systems for Video Technology 32.8 (2021): 5695-5706.
[16] Singh, Himanshu Kumar, and Amit Kumar Singh. “Comprehensive review of watermarking techniques in deep-learning environments.” Journal of Electronic Imaging 32.3 (2022): 031804.
[17] Kandi, Haribabu, Deepak Mishra, and Subrahmanyam RK Sai Gorthi. “Exploring the learning capabilities of convolutional neural networks for robust image watermarking.” Computers & Security 65 (2017): 247-268.
[18] Wu, Jinlin, et al. “De-Mark GAN: Removing dense watermark with generative adversarial network.” 2018 International Conference on Biometrics (ICB). IEEE, 2018.
[19] Bagheri, Mahnoosh, et al. “Adaptive control of embedding strength in image watermarking using neural networks.” arXiv preprint arXiv:2001.03251(2020).
[20] Zheng, Wenbo, et al. “Robust and high capacity watermarking for image based on DWT-SVD and CNN.” 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA). IEEE, 2018.
[21] Mellimi, Sandeep, et al. “A fast and efficient image watermarking scheme based on deep neural network.” Pattern Recognition Letters 151 (2021): 222-228.
[22] Islam, Mohiul, Amarjit Roy, and Rabul Hussain Laskar. “SVM-based robust image watermarking technique in LWT domain using different sub-bands.” Neural Computing and Applications 32 (2020): 1379-1403.
[23] Kang, Xiaobing, et al. “Multi-dimensional particle swarm optimization for robust blind image watermarking using intertwining logistic map and hybrid domain.” Soft Computing 24 (2020): 10561-10584.
[24] Ge, S., et al. “A Robust Document Image Watermarking Scheme using Deep Neural Network. arXiv 2022.” arXiv preprint arXiv:2202.13067.
[25] Zhu, Jiren, et al. “Hidden: Hiding data with deep networks.” Proceedings of the European conference on computer vision (ECCV). 2018.
[26] Liu, Yang, et al. “A novel two-stage separable deep learning framework for practical blind watermarking.” Proceedings of the 27th ACM International conference on multimedia. 2019.
[27] Tancik, Matthew, Ben Mildenhall, and Ren Ng. “Stegastamp: Invisible hyperlinks in physical photographs.” Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020.
[28] Mahapatra, Debolina, et al. “Autoencoder-convolutional neural network-based embedding and extraction model for image watermarking.” Journal of Electronic Imaging 32.2 (2022): 021604.
[29] Rahim, Rafia, and Shahroz Nadeem. “End-to-end trained cnn encoder-decoder networks for image steganography.” Proceedings of the European Conference on Computer Vision (ECCV) Workshops. 2018.
[30] Ding, Weiping, et al. “A generalized deep neural network approach for digital watermarking analysis.” IEEE Transactions on Emerging Topics in Computational Intelligence 6.3 (2021): 613-627.
[31] Kaczyski, Maciej, Zbigniew Piotrowski, and Dymitr Pietrow. “High-Quality Video Watermarking Based on Deep Neural Networks for Video with HEVC Compression.” Sensors 22.19 (2022): 7552.
[32] Fei, Jianwei, et al. “Supervised gan watermarking for intellectual property protection.” 2022 IEEE International Workshop on Information Forensics and Security (WIFS). IEEE, 2022.
[33] Yu, Ning, et al. “Artificial fingerprinting for generative models: Rooting deep fake attribution in training data.” Proceedings of the IEEE/CVF International conference on computer vision. 2021.
[34] Li, Daming, et al. “A novel CNN based security guaranteed image watermarking generation scenario for smart city applications.” Information Sciences 479 (2019): 432-447.