Investigating the Effectiveness of Artificial Intelligence in Watermarking and Steganography for Digital Media Security

Shadan M.J.a Abdalwahid
Information System Engineering Dept.
Engineering College
Erbil Polytechnic University
Erbil, Iraq
shadan.abdalwahid@epu.edu.iq

Wassan Adnan Hashim
Medical Instruments techniques Dept
AlQAlam University College
Kirkuk, Iraq
wasan.eng@alqalam.edu.iq

Mohammed Ganim Saeed
Department of Information Technology
Dohuk Polytechnic University
Dohuk, Iraq
Mohammed.saeed@dpu.edu.krd

Sarmad A. Altaie
Computer Engineering Department
University of Technology
Baghdad, Iraq
sarmad.a.altaie@uotechnology.edu.iq

Shahab Wahhab Kareem
Information System Engineering Dept.
Erbil Technical Engineering College
Erbil Polytechnic University
Erbil, Iraq
shahab.kareem@epu.edu.iq

2024 21st International Multi-Conference on Systems, Signals &amp; Devices (SSD) | 979-8-3503-7413-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/SSD61670.2024.10549272

Abstract— **Watermarking and Steganography are methods of embedding digital information within images or other media, such as text or audio, for the purpose of Copyright Protection or covert communication. This field of study is not recent and has been ongoing for several years, culminating in its current advanced stage. The utilization of Artificial Intelligence algorithms has played a pivotal role in revolutionizing various aspects, including security concerns and the precision of outcomes, as compared to traditional methods. This paper focuses on doing an in-depth analysis of cutting-edge research, techniques, and methodologies employed in the domain of Watermarking and Steganography, specifically in conjunction with Artificial Intelligence. By thoroughly examining and evaluating a collection of recent studies in this domain, we have scrutinized the outcomes of each study with respect to its research goal, the acquired results, the employed algorithm, and the research's robustness in terms of susceptibility to various forms of attacks and the technique of data embedding. Our findings indicate that the use of Artificial Intelligence algorithms has a substantial influence on enhancing result precision, system resilience, and establishing data ownership. Deep Neural Networks (DNN) are essential in Watermarking and Steganography due to their robustness, effectiveness, and accuracy. Novel methodologies and systems improve security, incorporation rates, precision of detection, and speed of convergence. Deep learning is being investigated in techniques such as data concealment, information hiding, and steganography to improve security.**

Keyword- Watermarking, Steganography, Artificial intelligence, Key Embedding

I. INTRODUCTION (*HEADING 1*)

The quick growth of digital tools has taken us into the world of computers and technology[3], this change has impacted different parts of our lives, including how we talk and enjoy fun things, education, health and more. Today, keeping digital data safe is very important, using ways like stealth coding and marking things digitally is very important for making sure that, Protection of different types of multimedia information, like pictures, sound recordings and video files. The way we hide information using these methods is constantly changing and it plays an important role play a big part in making digital content safe[4] [1] [5].

In recent years, the field of watermarking and steganography with artificial intelligent(AI) has witnessed significant advancements in techniques for embedding information (watermarks) into digital media to protect intellectual property, applications in image, audio, and video watermarking, use of AI for robust and imperceptible watermarking, concealing information within digital media to ensure covert. communication, techniques for hiding data within images, audio, and other media, integration of AI for enhanced hiding and detection mechanisms. The preservation of all digital media during transfer and storage, as well as the safe and secure preservation of texts so that they are undetectable to third parties, are all significant issues.

This review of article addresses the issue by conducting a comprehensive analysis of contemporary technologies employed by a group of researchers. They propose a novel approach to incorporating, extracting, and concealing data within digital information. The article also explores the advantages of utilizing artificial intelligence in this domain and provides guidance on selecting the most suitable algorithms to enhance system efficiency. Furthermore, by pinpointing the most crucial approaches that enhance the system's effectiveness, researchers can subsequently concentrate on refining these methods and ensuring their up-to-date implementation. It also helps a new researcher in the field choose the best technology and data set for a specific application by knowing the benefits and disadvantages of each technology. Hence, the issue is referred to as watermarking and steganography. The subsequent parts of the review paper are structured as follows:- part-2-Related Work. Part-3-focuses on digital media. Part-4-presents the results and discussion. Part-5-provides a Conclusion summary of the review.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

II. RELATED WORKS

In [6] The proposed method introduces digital watermark techniques for deep neural networks, offering a versatile frame work for integrating diverse watermarks into networks. This method allows remote verification of model ownership through embedded watermarks, allowing transparent and opaque access. The framework also enables remote ownership authentication with minimal API requests. analysis on two datasets show that its compliance with watermarking standards and resilience against most of watermark attacks, demonstrating its potential for enhance the security of deep neural network . The authors of article [7], proposed a modern steganography techniques that encode the maximum msg bits within edge pixels by using edge detection. The technique gradually decreases the number of encoded bits as the image transitions from distinct features to uniform characteristics, aiming for low detectability by the human visual system. The method uses a CNN algorithm and a Deep Supervision-based edge detector to evaluate edges in cover images. It reduces thresholding techniques during the conversion of binary and outperforms existing strategies in terms of peak signal-to-noise ratio and payload capacity in spatial and edge based steganography. In article [8], the authors protect intellectual property rights by presenting a comprehensive methodology for integrating watermarks into deep neural network models. The authors address the challenges of incorporating watermarks into deep neural networks, specifying important conditions, embedding techniques, and various types of attacks. The researcher propose a parameter regularizer methodology to ensure network performance remains unaffected by the embedded watermark. collection of tests explain the effectiveness of embedding watermarks in deep neural networks, with the frame work retaining the embedded watermark even after removing 65% of parameters. The authors in [9], The study provided a complete comprehensives deep learning techniques in steganalysis, concentrate on feature learning and updating a new paradigm. It introduces a feature learning technique that incorporates global information constraints, by focusing in the importance of global information in expressing steganalysis features. The method's detection performance appears the effectively comparable with previous CNN-based approaches, offering a modern perspective on steganalysis that effectively catch the statistical attributes of steganographic images. The article holds promising implications for steganography of image and deep learning. The embedding system presented in [10] the authors combined the term of digital watermarking to deep neural networks, the target to remotely verify the real ownership of DNN models from embedded watermarks. The researchers explored different watermarking frame works, integrated them into DNN models, and conducted ownership verification. The DNN model show very good successful owner verification depending on the embedded watermarks. Performance evaluations were carried out on standard datasets, the result showing that our frame work aligns with general watermarking standards and exhibits robustness vs the potential attacks.

The authors of paper [11] evaluated the effectiveness of detection hidden information in JPEG images using shallow ensemble classifiers with deep learning techniques. from the results appear clearly that performance varied depending on the steganographic technique and the level of hidden data encoded. The NSF5 algorithm was highly accurate in discovering content concealed, but J-Unwired was difficult to discover. The system evaluation aimed to calculate the most effective feature space for image steganalysis. The DCTR and GFR parameters yielded better results than PHARM parameters, suggesting the use of these properties during security checks of JPEG files. Also the top performing deep learning technique was either on par or slightly less efficient than the leading ensemble classifier constructed using linear regression. These outcomes suggest that carefully chosen ensemble classifiers could be a good alternative for deep learning techniques in image steganalysis. The method presented in [12] explanation of a new practical Two-Stage separable Deep Learning (TSDL) frame work that has been specifically developed for blind watermarking types.The architecture comprises of two stages: noise-free end-to-end adversary training (FEAT) and noise-aware decoder-only training (ADOT). Our thorough trials confirm that the TSDL framework is resilient to both typical classical high-intensity noises and certain black-box noises that prior studies did not evaluate. Our TSDL framework outperforms state-of-the-art approaches in terms of performance, regardless of the type of noise. Furthermore, the system seeks to elucidate the mechanism of blind watermarking using deep learning. We posit that this comprehension will facilitate additional progress in this technology and serve as a source of motivation for future research endeavors. The research [13] proposed system introduce a novel approach to enhance the security of deep learning-based steganography methods. Leveraging the linear behavior of state-of-the-art CNN-based steganalyzers, the system employ adversarial example techniques to enable steganographic content to evade detection. The distortions introduced to stego content by the adversarial technique are evaluated using Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR). Another point recognize the limitations of traditional indicators like MSE and PSNR in detecting localized distortions and express the need for further research to minimize the impact of adversarial perturbations on the secrecy of information. Additionally, the proposed investigating post-silicon technologies, such as nanofluidics and microfluidics devices, to address security concerns in novel hardware systems. Another article of watermarking [14] the authors present a novel approach to deep blind watermarking, introducing a new encoder design for deep blind watermarking models and a watermarking strategy that addresses the balance between robustness and imperceptibility. The proposed method achieves an average 5.46 higher Peak Signal-to-Noise Ratio (PSNR) than baseline methods, ensuring comparable robustness. Additionally, the deep blind watermarking strategy reduces training time and Video Random Access Memory (VRAM) usage by less than 1/4 when encoding/decoding in a 3 × 256 × 256 image. Given that many real-world images are likely larger than 3 × 256 × 256, our method is expected to offer even greater.

performance gains in terms of training time and computational resources. As a result, proposed method effectively addresses the trade-off in deep blind watermarking with lightweight training. The paper [15] a time-series steganography method is introduced, employing adjacent mean values to embed secrets within a cover image. With a bit per pixel rate of 7.4, the PSNR is 31.26; and with a bit per pixel rate of 8.88, the PSNR is 25.77. The proposed method demonstrates strong performance across various metrics, including SSIM, Correlation, Intersection, and LSB enhancement. The authors [16] Presents a new design that incorporates a dual attention mechanism and simulated JPEG compression for blind picture watermarking. The suggested model achieves an improved equilibrium between resilience and invisibility by incorporating channel attention and spatial attention into the network architecture. The results of many testing unequivocally establish the superiority of this technology compared to earlier approaches, in terms of both image quality and robustness. In study [17] The research presents the Reversible-Logic-Based-Hexel-Value-Differencing (RLBHVD) technique in the context of “Hexel Interpolation Prediction” (HIP) domain. The technique is specifically designed to hide Spatial-Domain (S-D) data. The efficacy of a data concealment technique relies on its capacity to conceal data within the content while minimising the degree of modification between the original and modified copies resulting from the data concealment process. Simulations demonstrate that steganography in the HIP domain outperforms steganography in the SIP domain in terms of data concealment capacity and the degree of modification imposed on the material during data concealing. Article [18] the authors introduces a Multi Task Learning (MTL) based Deep Neural Network (DNN) watermarking model for ownership verification. The authors outline the fundamental security requirements for DNN watermarking and address privacy concerns. Their proposal involves embedding a watermark as an additional task alongside the primary task. The scheme incorporates regularizers to explicitly meet various security requirements. The design of the watermarking task, along with these regularizers, ensures tractable security for the MTL-based DNN watermarking scheme. Additionally, a decentralized consensus protocol is employed to make the entire framework secure against potential attacks. Paper [19] The article discusses uncharted topics, including the effect of intensity initialization on predictive accuracy and the role of distributional shift in dual-layer prediction. Empirical findings demonstrate that the act of assigning a pixel's intensity to zero, although it may seem random, consistently results in a minimal loss across multiple epochs. Moreover, training models in a causal manner might somewhat reduce distributional shift in deployment by minimising the difference in distribution between training and test sets. State-of-the-art prediction accuracy and steganographic rate-distortion performance can be achieved using advanced pixel-level computer vision models. The authors of work [20] DLWIoT is a framework specifically developed for facilitating the integration of IoT devices through the utilization of image watermarking techniques. DLWIoT enables the covert integration of user credentials within an image, such as a QR code, that is physically imprinted on an IoT device. The embedded information can be exclusively utilized by an authorized user for device onboarding. The system utilizes an innovative deep learning-based method for embedding watermarks in photos, which showcases resilience against distortion and degradation in the quality of the tagged images. The article [21], a novel approach is proposed for concealing confidential information within cover text using artificial intelligence and deep learning theories, specifically employing the Long Short-Term Memory (LSTM) theory. The generation of sentences and texts carrying confidential information is achieved through LSTM. Unlike previous methods, the hiding process operates at the letter level rather than the word level, enhancing the cover text's capacity to carry more hidden bits. This approach generates multiple texts simultaneously, allowing for the selection of the most suitable texts for a given context. The study successfully achieves the desired goal in text steganography. The researchers in [22] introduce an innovative technique of digital watermarking specifically for insert binary sequences into anonymized data. The target of this technique is to discourage unauthorized, illegal redistributions on platforms and infrastructure that enable the spread of anonymized data. By integrating digital watermarks into the publication system, it becomes feasible to authenticate the source of an unauthorized redistribution, clearly identifying the ownership of the content. By comparing the evaluation findings with the previous methodology, it was shown that our system can encode bit strings with minimum information loss and strong resistance to distortion attacks. The new technique demonstrated reduced levels of information loss in comparison to the prior one. In [23], The researchers developed a simple and effective frame work that allows for the hiding of information within digital image from the entire process. The suggested system comprises an extraction network, embedding network, and preprocessing module, which is influenced by the deep convolutional of the auto encoder. The preprocessing module prepares input images for the embedding network to hide the secret image within the cover image, while the extraction network retrieves the secret image from the stego image generated by the embedding network. The suggested technique showing higher PSNR values, indicating enhanced security and robustness compared to classical and deep learning image steganography algorithms. Also the proposed method exhibited outstanding camouflage, producing steganographic images that very close matched the original cover image. The work [24] proposes a method to combine watermarks with deep neural network (DNN) hardware accelerators to prevent intellectual property theft. The article techniques, based on p-ADMM, optimizes hardware overhead and watermark impact on the design's algorithmic functionality. The study found minimal alterations in performance on picture classification models, also found minimal overall changes if a lesser Equivalent Series Resistance (ESR) is acceptable. The addition of a watermark to a DNN accelerator incurs minimal hardware overhead, resulting in a 0.18% increase in Look-Up Tables (LUTs) usage and a 0.17% increase in power consumption. This technique effectively protects vs IP infringement while

Governorate the original functionality of the hardware architecture. The research paper in [25] Presented a hybrid deep learning model (HDLM)system for reversible image steganography, can be divided to three models: one- a convolutional neural network (CNN), two-a cycle-consistent generative adversarial network (CycleGAN) for encoding, and three-a deep neural network (DNN) for cover selection. The security test included StegoExpose, while the HDLM method was evaluated using metrics such as payload capacity, PSNR, and Structural Similarity Index (SSIM), Based on the results, the HDLM algorithm outperformed other deep learning models in terms of SSIM values, PSNR, and payload capacity. The study's findings have significant implications for the domains of technology and computer science, especially in the area of image steganography.

III. DIGITAL MEDIA

A ubiquitous technological requirement in the smart life is multimedia information processing, particularly for digital media optimization applications. For this field, dependable visual data processing systems with a mature design are still lacking. Due to the high dynamics and volume of commerce associated with digital media contact, there is a greater need for processing power and efficiency[26].

A. Digital Watermarks

value at the terminals of the generator bus during steady state Digital watermarks are very important in keeping ownership safe, making sure it's real and checked, adding notes or comments to different online stuff. A good watermarking system should be able to show strong resistance to many kinds of changes, while typical image editing, like smoothing out or squashing, can affect the watermark's strength by changing picture values and shape distortions pose an additional challenge. Geometric distortions disrupt synchronization between the watermark writer and reader, making it necessary to fix these problems for a strong one watermarking system [27] [28] [29].

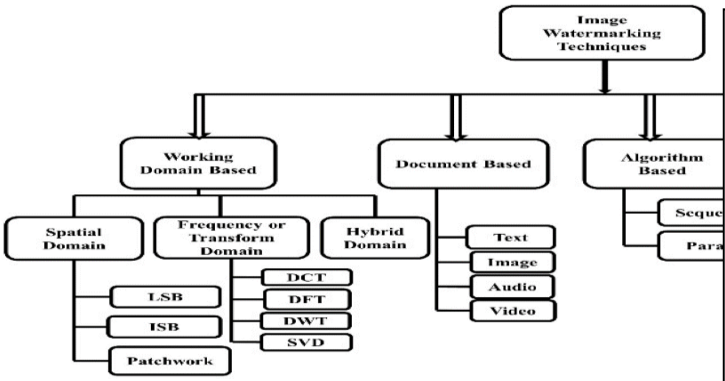

Fig. 1. Watermarking tec

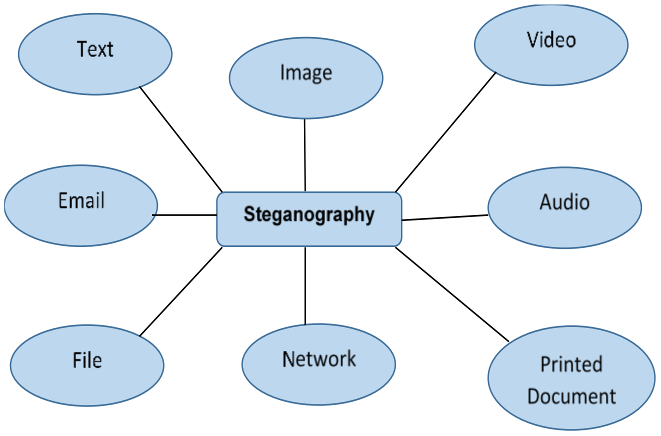

Fig. 2. Steganography Techniques

The procedure for watermarking approaches is incorporating a watermark—a marker of the owner's authenticity—into a host signal. It is possible to retrieve this watermark data afterward. The watermark can have different characteristics, such as being visible or invisible, and it might be one bit, a set of binary data, or many samples from the host signal. This technique is often employed for purposes such as ownership verification, authentication, and content protection Figure 1 show some techniques [30].

B. Digital Stegangraphy

Steganography is the act of hiding a confidential communication within a host medium, aiming to avoid detection by any monitoring party, commonly referred to as the "warden." Digital image data is a prevalent medium for steganography due to its high redundancy, providing ample capacity for embedding. In the early stages of digital image steganography, both spatial and compressed domain approaches were non-adaptive, treating all pixels or discrete cosine transform (DCT) coefficients uniformly, Steganography can be classified [31] [32] [33]. Steganography involves the covert transmission of messages or information to a recipient, concealed within various entities like letters, photographs, documents, and more. The term "Steganography" is derived from the Greek terms "Stegos," which means "to cover," and "graya," meaning "writing" or "books," essentially translating to "secret

Steganography is widely employed to enhance security and ensure data safety. The primary objective of Steganography

Steganography and watermarking, although related, serve distinct purposes, and the terms are sometimes used interchangeably[34]. Watermarking primarily aims to create a permanent mark on a message, making it difficult for an intruder to eliminate or substitute the message. This technique is commonly employed for asserting ownership, ensuring authenticity, or encoding Digital Rights Management (DRM) rules[35]. writing." In the context of internet data transfer, The message embedded through watermarking may or may not be invisible. On the other hand, the main objective of steganography is to facilitate one-to-one communication while concealing the existence of a message.

In the context of steganography, an intruder, who lacks knowledge of what to look for, cannot even detect the presence of a message within the data [36] [37]. Table 1 showing the prosperities of Watermarking and Steganography.

C. Artificial Intellegint

The media has paid considerable attention to the topic of artificial intelligence. Certain nations assert to be at the forefront of their respective fields, while others assert to be emerging victorious in the competition for artificial intelligence leadership. Since ancient times, artificial intelligence has been a matter of discussion in our society[38] [39]. Education plays a crucial role in the evolution of civilization, and its approaches, subjects covered, ideas, and role models are all evolving. The idea that educational problems can be solved by utilizing artificial intelligence's (AI) potential has gained popularity recently. The state of the art for integrating AI in K–12 education was presented in this study. In particular, several K–12 grades and courses were taken into consideration while discussing the various areas of education where AI was used along with the associated AI categories [40]. Artificial intelligence has experienced rapid growth in recent years as computers and servers have become increasingly adept at simulating human performance. Many issues with natural language processing, image recognition, and audio recognition have been resolved thanks to artificial intelligence. A digital watermark added to the original versions of documents exchanged over the internet has also assisted in bolstering their security. Machine learning techniques have been used recently to guarantee the security of data transferred via the internet. Machine learning techniques in watermarking algorithms can be used to speed up the training set, predict the ideal embedding strength, strike a compromise between resilience and security, build and optimize the maximum likelihood relation set and more [41]. The secret image is concealed into the cover image using the hiding network, which employs a neural network topology based on the Swin-Transformer. In Figure 3, the particular structure is displayed [42].

D. Performance of Key Embedding

In order to analyze watermarked images and extracted watermarks, performance analysis is essential. As seen in figure (4), many statistical indicators are employed to examine performance. There is a visual degradation of the image because the watermark robustness is only dependent on the strength of the watermark embedding. These visual deteriorations are useful for assessing performance [43].Authors frequently take key performance indicators into account when measuring the quality of developed algorithms in order to determine their efficacy. A crucial evaluation criterion is the notion of invisibility, especially for algorithms intended to hide information in digital images (apart from some watermark embedding systems). Under such circumstances, the information that is embedded ought to remain invisible to the naked sight [1].

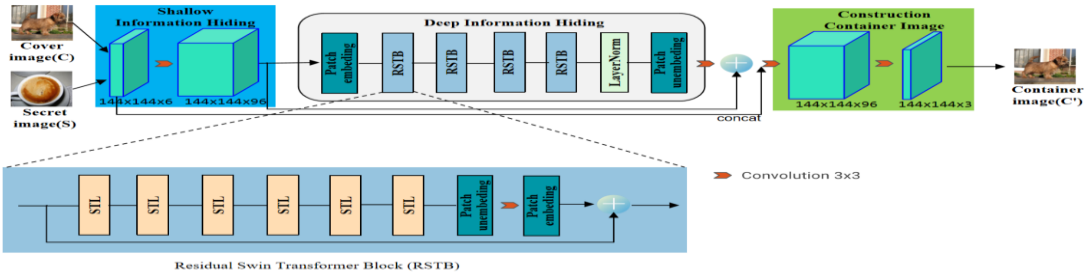

Fig. 3. Architecture of the artificial hiding network.

For quantitative invisibility estimation, peak signal-to-noise ratio (PSNR) is typically used in investigations, additionally measured the quality of reconstruction of lossy image compression codecs using quality assessment technique[44]. The following formula is used to compute the PSNR when comparing an embedded image and a container image:

$$\text{PSNR (dB)} = 10 \times \log_{10}\left(\frac{255^2}{\text{MSE}}\right), \qquad (1)$$

TABLE 1. PARAMETERS OF WATERMARKING AND STEGANOGRAPHY[1, 2]

Parameter	Watermarking	Steganography
Definition	The practice science of hiding information	Process of hiding information-in carrier signal or hardware
Used for	Control of integrity, authentication, ownership	Hidden data transmission
Protection	Original image	secret data
Input data	Image, video, text	All types of digital data
Output data	Watermarked Data	Stego Data
Capacity Require	Medium	High
Extract embedding Information	Depend on cases	Yes
Security Attack	Image distortion	Steganalysis
Visibility	Some time	Never

the two objectives of the Watermarking task are to minimize cover distortion and maximize the recovery of watermark information, The watermark and cover images' distortion is measured using the mean squared error, or MSE [45]. Equation of MSE as shown below:

$$\text{MSE}=\frac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N}(pij-qij)^2 \qquad (2)$$

where p is the cover of image pixel and q are the stege image. The M and N denote the image's height, and widths [46]

The Structural Similarity Index Method(SSIM) Perception-based technique is deployed. This approach considers image deterioration as a change in the perception of structural information. Additionally, it operates with several other essential perception-related factors, such as contrast and brightness masking. The term "structural information" refers to pixels that are either spatially limited or strongly interconnected [44] [47] . The formula of SSIM as shown in equation 3:

$$SSIM = \frac{(2\mu_C\mu_S + K_1)\times(2\sigma_{CS} + K_2)}{(\mu_C^2 + \mu_S^2 + K_1)\times(\sigma_C^2 + \sigma_S^2 + K_2)}, \qquad (3)$$

Where σ2c = Pixel value variance of the cover image , σ2s= Variance of pixel values in the in the attached image. , σcs= covariance of both pictures, K1 and K2 are constants, and μc = Mean pixel value of the cover image and μs = the mean pixel value of the image provided as an attachment[1]

E. Watermarking and steganography application

Data hiding methods have been widely utilized to provide copyright safeguarding, data consistency, clandestine communication, non-denial, and verification, among numerous other uses [48] [49].

Data concealment techniques offer a wide range of possible uses. Digital watermarking has various applications, such as protecting copyright by providing evidence of ownership, identifying owners, or tracking transactions. It is also used for monitoring broadcasts, authenticating content by detecting tampering or identifying its location, controlling copies, managing devices, and enhancing legacy systems. Steganography, on the other hand, focuses on hidden communication, specifically for the use of military operations, dissident activities, or criminal organizations. The utilization of steganography for military and criminal purposes has generated a rising fascination among scholars in the field of steganalysis, which pertains to the techniques employed to identify concealed communications. [50] [2].

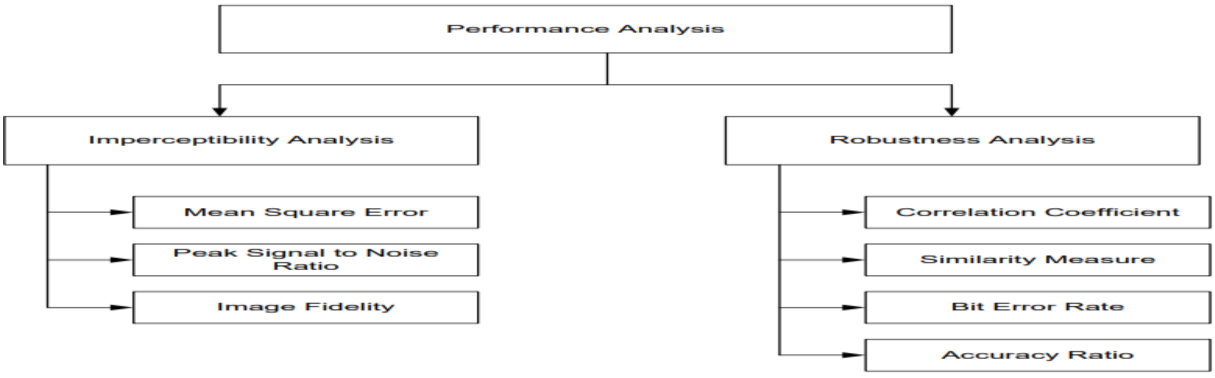

Fig. 4. Metrics for watermarking technology performance analysis.

In recent years, new application situations have evolved. The novel uses of data hiding include protecting privacy during transaction tracking, establishing data provenance through digital watermarking, aiding forensic investigations through digital watermarking, concealing information within network flows, hiding information within network communications, concealing information within VoIP communications, enabling criminals and terrorists to hide information, and facilitating malware injection through hidden channels, among other applications. [51].

IV. RESULTS AND DISCUSSION

The models chosen for this study were selected based on past research and studies indicated in the literature review section. Samples from the results were then picked to clarify and compare these models, as detailed in the table 2. Our table showing the summary analysis of different techniques and methods that relate in fusion digital media, watermarking and steganography in various fields including transform domain, spatial domain, DCT, DNN are also used in the domains of digital media, showing how they play as good role in both of water and stego. the analysis of most of articles point out that steganography and watermarking must achieve flexibility, optimal efficiency and very good accuracy. the datasets for most articles are MNIST, CIFAR 10, CIFAR 101, COCO, Boss base, LFW, USC SIPI and others. COCO is most frequently used, sometimes CIFAR. There are several performance metrics used to evaluate the suggested methods including PSNR, MSE, SSIM, AUC and ROC. The summary also appears that the algorithm deep learning work as good classifiers in steganography detection as well extract several feature spaces for the purpose of detection. In essence, this study article researches the different types to confirm the real owner of the digital media, intellectual property protection and outside ownership verification of deep learning models. the summary of the analysis presents original solutions, models, and methods to improve security and embedding rates, also the accuracy of detection for the watermarking design and steganography. There are several publications that Reach out into the use of deep learning in data concealment, information hiding and making steganography techniques more secure. The discussed methodologies and algorithms demonstrate good results in field of security, robustness vs different types of attacks and distortions are concerned and the ability to hide data.

V. CONCLUSION

Steganography is the technique employed to covertly convey secret information by concealing it within the visible image of the cover, and watermarking refers to the incorporation of a code or digital image, which can be either visible or invisible, into multimedia material. At present, there is significant progress in the development of differential steganography methods and digital watermarks. This progress is primarily driven by important scientific advancements in artificial intelligence especially deep learning algorithms, which include the integration of several algorithms. Researchers from many countries provide numerous novel algorithms that vary based on quality characteristics. from our article most researchers have shown the efficiency of DNN algorithm in this field with the different types of datasets used e.g. COCO and CIFAR, also most researchers used robustness measurement standards that shows the power of the system to withstand various attacks and common distortions, and others lacked it. Although there is a wide range of advanced algorithms, there are still several unresolved issues in the field of information that are present in digital watermark and steganography. The use of artificial intelligence technologies has greatly contributed to the development of the system in this field. The review indicated that efforts were being made in these domains, however numerous challenges persist that necessitate the development of novel local remedies.

References

[1] Awla, H.Q., A.R. Mirza, and S.W. Kareem. An Automated CAPTCHA for Website Protection Based on User Behavioral Model. in 2022 8th International Engineering Conference on Sustainable Technology and Development (IEC). 2022. IEEE.

[2] Ting, H.L.J., et al., On the trust and trust modeling for the future fully-connected digital world: A comprehensive study. IEEE Access, 2021. **9**: p. 106743-106783.

[3] Evsutin, O., A. Melman, and R. Meshcheryakov, Digital steganography and watermarking for digital images: A review of current research directions. IEEE Access, 2020. **8**: p. 166589-166611.

[4] Yousif, M.K., Z.E. Dallalbashi, and S.W. Kareem, Information security for big data using the NTRUEncrypt method. Measurement: Sensors, 2023. **27**: p. 100738.

[5] Zhang, J., et al. Protecting intellectual property of deep neural networks with watermarking. in Proceedings of the 2018 on Asia conference on computer and communications security. 2018.

[6] Ray, B., et al., Image steganography using deep learning based edge detection. Multimedia Tools and Applications, 2021. **80**(24): p. 33475-33503.

[7] Nagai, Y., et al., Digital watermarking for deep neural networks. International Journal of Multimedia Information Retrieval, 2018. **7**: p. 3-16.

[8] Zou, Y., G. Zhang, and L. Liu, Research on image steganography analysis based on deep learning. Journal of Visual Communication and Image Representation, 2019. **60**: p. 266-275.

[9] Deeba, F., et al., Digital watermarking using deep neural network. International Journal of Machine Learning and Computing, 2020. **10**(2): p. 277-282.

[10] Płachta, M., et al., Detection of image steganography using deep learning and ensemble classifiers. Electronics, 2022. **11**(10): p. 1565.

[11] Liu, Y., et al. A novel two-stage separable deep learning framework for practical blind watermarking. in Proceedings of the 27th ACM International conference on multimedia. 2019.

[12] Shang, Y., et al., Enhancing the security of deep learning steganography via adversarial examples. Mathematics, 2020. **8**(9): p. 1446.

[13] Yang, B., G. Lim, and J. Hur, Toward Practical Deep Blind Watermarking for Traitor Tracing. IEEE Access, 2023.

[14] Chuang, Y.-H., et al., Steganography in RGB images using adjacent mean. IEEE access, 2021. 9: p. 164256-164274.

[15] Zhong, J.-Y., et al., Enhanced Attention Mechanism-Based Image Watermarking With Simulated JPEG Compression. IEEE Access, 2023. **11**: p. 135934-135943.

[16] Cevik, T., et al., Reversible Logic-Based Hexel Value Differencing—A Spatial Domain Steganography Method for Hexagonal Image Processing. IEEE Access, 2023. **11**: p. 118186-118203.

[17] Li, F. and S. Wang, Secure watermark for deep neural networks with multi-task learning. arXiv preprint arXiv:2103.10021, 2021.

[18] Chang, C.-C., et al., Deep learning for predictive analytics in reversible steganography. IEEE Access, 2023. **11**: p. 3494-3510.

[19] Mastorakis, S., et al. Dlwiot: Deep learning-based watermarking for authorized iot onboarding. in 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC). 2021. IEEE.

[20] Adeeb, O.F.A. and S.J. Kabudian, Arabic text steganography based on deep learning methods. IEEE Access, 2022. **10**: p. 94403-94416.

[21] Nakamura, Y. and H. Nishi, Digital Watermarking for Anonymized Data With Low Information Loss. IEEE Access, 2021. **9**: p. 130570-130585.

[22] Subramanian, N., et al., End-to-end image steganography using deep convolutional autoencoders. IEEE Access, 2021. **9**: p. 135585-135593..

[23] Clements, J. and Y. Lao. DeepHardMark: Towards watermarking neural network hardware. in Proceedings of the AAAI Conference on Artificial Intelligence. 2022.

[24] Oludele, A., et al., Security test using StegoExpose on hybrid deep learning model for reversible image steganography. Computer Science Department, Babcock University, Ilishan-Remo, Ogun State, Nigeria, 2022.

[25] Lei, X., Design of A Deep Neural Network-based Visual Data Processing System for Digital Media Optimization Applications. IEEE Access, 2023.

[26] Ma, Z., et al., Local geometric distortions resilient watermarking scheme based on symmetry. IEEE Transactions on Circuits and Systems for Video Technology, 2021. **31**(12): p. 4826-4839.

[27] Khalifa, F.M. and M.G. Saeed, Image Watermarking Using All Phase Discrete Cosine Biorthogonal Transform in Selected Pixel Blocks. Polytechnic Journal, 2020. **10**(1): p. 68-73..

[28] Zhang, Y.-Q., et al., DeepTrigger: A watermarking scheme of deep learning models based on chaotic automatic data annotation. IEEE Access, 2020. **8**: p. 213296-213305.

[29] Pushpa Mala, S., D. Jayadevappa, and K. Ezhilarasan, Digital image watermarking techniques: a review. Int J Comput Sci Secur, 2015. **9**(3): p. 140-156.

[30] Aloraini, M., M. Sharifzadeh, and D. Schonfeld, Quantized Gaussian JPEG steganography and pool steganalysis. IEEE Access, 2022. **10**: p. 38031-38044.

[31] Su, W., et al., A new distortion function design for JPEG steganography using the generalized uniform embedding strategy. IEEE Transactions on Circuits and Systems for Video Technology, 2018. **28**(12): p. 3545-3549.

[32] Rahman, S., et al., A novel steganography technique for digital images using the least significant bit substitution method. IEEE Access, 2022. **10**: p. 124053-124075.

[33] Gurunath, R., et al., A novel approach for linguistic steganography evaluation based on artificial neural networks. IEEE Access, 2021. **9**: p. 120869-120879.

[34] Abuali, M.S., et al., Digital image steganography in spatial domain a comprehensive review. J. Theoretical Appl. Inf. Technol, 2019. **97**(19): p. 5081-5102.

[35] Azade Abedini, Ghazanfar Shahgholian, and Bahdor Fani, "Power Mikhail, D.Y., R.S. Hawezi, and S.W. Kareem, An Ensemble Transfer Learning Model for Detecting Stego Images. Applied Sciences, 2023. **13**(12): p. 7021.

[36] Tanha, M., et al. An overview of attacks against digital watermarking and their respective countermeasures. in Proceedings Title: 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec). 2012. IEEE.

[37] Arsenova, E. Technical aspects of digital rights management. in Seminar: Digital Rights Management. 2008.

[38] Ghanem, S., Digital Rights Management of Image Using LSB Embedding Lucas Sequence. 2023.

[39] Tao, H., et al., Robust image watermarking theories and techniques: A review. Journal of applied research and technology, 2014. **12**(1): p. 122-138..

[40] Ramakrishnan, S., Digital Image and Video Watermarking and Steganography. 2019: IntechOpen

[41] Radanliev, P., et al., Forecasts on future evolution of artificial intelligence and intelligent systems. IEEE Access, 2022. **10**: p. 45280-45288

[42] Hawezi, R.S., F.S. Khoshaba, and S.W. Kareem, A comparison of automated classification techniques for image processing in video internet of things. Computers and Electrical Engineering, 2022. **101**: p. 108074

[43] Zafari, M., et al., Artificial intelligence applications in K-12 education: A systematic literature review. IEEE Access, 2022. **10**: p. 61905-61921

[44] El-den, B. and M.M. Eid, Watermarking Models and Artificial Intelligence

[45] Wang, Z., et al., Deep Image Steganography Using Transformer and Recursive Permutation. Entropy, 2022. **24**(7): p. 878

[46] Dixit, A. and R. Dixit, A review on digital image watermarking techniques. International Journal of Image, Graphics & Signal Processing, 2017. **9**(4): p. 56-66

[47] Sara, U., M. Akter, and M.S. Uddin, Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. Journal of Computer and Communications, 2019. **7**(3): p. 8-18

[48] Liao, X., J. Peng, and Y. Cao, GIFMarking: The robust watermarking for animated GIF based deep learning. Journal of Visual Communication and Image Representation, 2021. **79**: p. 103244

[49] Himthani, V., et al., Comparative performance assessment of deep learning based image steganography techniques. Scientific Reports, 2022. **12**(1): p. 16895

[50] Megías, D., W. Mazurczyk, and M. Kuribayashi, Data hiding and its applications: Digital watermarking and steganography. 2021, MDPI. p. 10928

[51] KAREEM, S.W., Secure Cloud Approach Based on Okamoto-Uchiyama Cryptosystem. Journal of Applied Computer Science & Mathematics, 2020. **14**(29)

[52] Abusham, E.A., et al., Fusion of Watermarking and Steganography for Protecting Image Ownership. Applied computing Journal, 2021: p. 152-164

[53] Megías, D. Data hiding: New opportunities for security and privacy? in Proceedings of the 2020 European Interdisciplinary Cybersecurity Conference. 2020

Table 2 watermarking and steganography data hiding

Ref.	purpose of use	Dataset	Domain	Embedding operation	Metrics	Algorithm	Robustness	Achieved	Aim and contribution
[6]	watermarking	MNIST and CIFAR10	transform domain	DNN	WM content, unrelated WM noise	DNN	high	The ownership is quickly, precisely, and 100% confirmed. It is strong and resilient.	Protecting the deep learning models' intellectual property and enabling external ownership verification
[7]	Steganography	Berkeley Segmentation	Spatial Domain	LSB	-	CNN	good	The proposed approach attains a greater payload.	The article expands upon the idea of "digital watermarking" by applying it to DNN models, in order to verify ownership and protect intellectual property.

[8]	watermarking	CIFAR-10 and Caltech-101	NND	embedding regularizer.	-	DNN	good	can incorporate a watermark when a deep neural network is being trained from the beginning.	Technology involving digital watermarking for authorizing ownership of deep neural networks.
[9]	Steganography	WOW and S-UNIWARD	DCT domain	adaptive algorithm	-	CNN	Medium	This study enhances steganalysis by utilizing advanced deep learning methods, emphasizing global information in feature learning, addressing low embedding rate challenges, and introducing a comprehensive detection strategy.	1. A novel steganalysis approach utilizing deep learning techniques. 2. A method for learning features in steganalysis. A technique for detecting hidden information with a low level of embedding in steganography using feature learning. 4. A comprehensive steganalysis technique for multi-class steganography
[10]	watermarking	Lena-image +image-set5	Spatial Domain	DNN		DNN	-	quickly and precisely confirms the ownership	Digital watermarking technology using DNN models
[11]	Steganography	BOSS	DCTR GFR PHARM	UERD	AUC ROC	LR	-	The nsF5 method achieved a density of 0.4 bpnzac with a 99.9% accuracy rate. the detection of J-Uniward at a density of 0.1 bpnzac was found to be extremely challenging, with a maximum accuracy rate of only 56.3%.	Image steganography detection Applying Deep Learning and Ensemble Classifiers to analyze various feature spaces for the purpose of detecting steganography.
[12]	watermarking	COCO	-	FEATADOT	PSNR	CNN	good	demonstrates improved performance, stability, and convergence speed as well as resistance to loud noises.	a brand-new, useful Two-stage Separable Deep Learning (TSDL) framework for watermarking blind images
[13]	Steganography	LFW, Bossbase	-	GANs	PSNR MSE	CNN SRM	medium	1) A new method for hiding information in a novel way, known as steganography scheme. 2) Improving the level of security. 3) Testing the effectiveness of the method through experiments. 4) Minimizing the amount of distortion	Present a novel steganography scheme that improves the security of deep learning Introduce a steganography scheme that takes use of the linear behavior of deep learning networks in higher-dimensional space and adds adversarial examples to boost security.
[14]	watermarking	COCO			PSNR	CFC+ CONCAT	medium	include watermarks into the deep learning hardware and run a ResNet ImageNet classifier with a 0:009% accuracy decrease.	adding watermarks to hardware accelerators for DNN.
[15]	Steganography	Lena +images	Time domain	adjacent mean	PSNR	-	-	The technique attains an embedding rate of around 7.4 bits per pixel (bpp) with a peak signal-to-noise ratio (PSNR) of nearly 30, and an embedding rate of about 8.88 bpp with a PSNR of almost 25.	To provide a new effective data concealment technique for 24-bit color photos. The contribution is the utilization of the spatial-domain-adjacent mean technique to conceal data within RGB images.
[16]	watermarking	COCO	DCT	-	PSNR SSIM	CNN	high	the model shows remarkable robustness against a range of common distortions and provides stronger tolerance to JPEG compression, with bit accuracy exceeding 99%.	A new improved attention-based technique for watermarking images that simulates JPEG compression
[17]	Steganography	USC–SIPI	Spatial-Domain	PVD	PSNR SSIM	RLB HVD	-	This study aimed to develop a new steganography method for hexagonal image processing, develop a software infrastructure, introduce heptad partitioning	Examine the utilization of steganography within the Hexel. Creates a tailored software framework to convert images into the HIP domain.

[18]	watermarking	MNIST CIFAR VirusShare	-	DNN	Rfunc RDA	DNN LSTM	medium	Because of its flexibility, security, and robustness, the suggested technique is a promising option for protecting deep learning models.	DNN watermarking methodology for ownership verification based on MTL.
[19]	Steganography	BOSSbase	-	pre-trained masked language model	PSNR SSIM	BERT	high	Comparative study of initialization strategies, Investigation of training strategies, Examination of loss functions, Assessment of predictive accuracy and steganographic performance.Illustration of the cutting-edge steganographic performance	Analyze the effects of various training setups on the predictive precision of neural networks and offer practical observations. Gain comprehension of deep learning models for predicting pixel intensity in reversible steganography.
[20]	watermarking	COCO CIFAR	automated system	binary-encoded	PSNR BER	DLWIoT	high	include data, such the user's voice (a recorded word or phrase, for example). Several biometrics, such as user movements, iris scans, and fingerprints.	introduced DLWIoT, a framework that uses image watermarking to facilitate IoT onboarding.
[21]	Steganography	Al Mutanabbi's ,Nizar Qabbani's poetry Words	-	LSTM	-	RNN LSTM	-	Introduces a novel method of concealing confidential information within Arabic poetry through the utilization of artificial intelligence (AI) and Long Short-Term Memory (LSTM) theories.	The authors employs deep learning and artificial intelligence to introduce a novel method for concealing sensitive data within Arabic poetry. It aims to increase storage capacity, improve linguistic accuracy.
[22]	watermarking	Pseudo Personal Information Generator.	-	binary encoded	FastText geographic-distance-based	-	good	The detection of the system success rate for the addition attack was just 40% of the attack ratio. The failure to detect an adding attack is caused by the inaccurate identification of a modified tuple within a subgroup.	techniques of digital watermarking that produce a correlation between data consumers and anonymized data, while ensuring little loss of information.
[23]	Steganography	COCO, CelebA aImageNet	DL	Auto encoder	PSNR MSE	CNN GAN	good	proposed an innovative method for hiding sensitive data within vireos images using a deep convolutional autoencoder structure.	The system introduces a simple and advanced deep convolutional autoencoder structure for concealing and retrieving confidential images from steganographic images.
[24]	watermarking	Cifar10 Cifar100 ImageNet	hardware	embedding a signature	ESR/ Δ Acc Δ Fid	DNNs `p-ADMM	high	A novel approach using a sophisticated blind model and watermarking mechanism improved PSNR by 5.46, maintaining robustness, and reducing training time and VRAM usage by < 0.25	The essay explores deep blind watermarking models, their impact on robustness and imperceptibility, and proposes a strategy to improve robustness without compromising imperceptibility, enhancing digital media copyright protection.
[25]	Steganography	COCO	-	Cycle GAN	PSNR and SSIM.	HDLM-CNN DNN-CGAN	-	The paper main achievement is the development , updating and evaluation of a “Hybrid Deep Learning Model”for reversible image steganography, which successfully addresses the concerns about payload capacity and security.	The aim of the paper is to explain the difficulties associated with payload capacity and security in digital image steganography systems, which is the primary focus. 1) Hybrid Deep Learning Model Proposal 2) Assessment of Payload Capacity and Security