

# A COMPREHENSIVE REVIEW OF WATERMARK DETECTION AND RECOGNITION TECHNIQUES: CHALLENGES, APPLICATIONS, AND FUTURE PERSPECTIVES

## Abstract

In terms of digital media security and copyright protection, watermark detection and recognition have become significant study fields. Identification and preservation of multimedia asset ownership are essential tasks in a period of broad digital content sharing. This review paper explores the developments achieved in this area by providing a thorough evaluation of watermark detection and recognition methods. The paper explores the fundamentals of watermarking and gives an overview of the several watermarking techniques used in image, video, and audio content. It talks about the difficulties that watermark detection systems encounter, such as robustness against attacks, effective processing, and resistance to unintended changes. The article also evaluates watermarking systems rigorously, highlighting any potential flaws that adversaries might try to exploit. It investigates the preventative measures put forth to boost the reliability of watermarking systems, hence enhancing the security of multimedia assets.

**Keywords:** watermar detection; watermark recognition; natural scene images; deep learning

## Authors

### Safnaz

Department of Computer Science  
University College Mangalore  
Mangalore, India 575 001  
safnaz.ibrahim97@gmail.com

### Bharathi Pilar

Department of Computer Science  
University College Mangalore  
Mangalore, India 575 001  
Bharathi.pilar@gmail.com

## I. INTRODUCTION

The rapid growth of digital media and its widespread accessibility has opened up new challenges concerning the protection of intellectual property rights and content ownership. To address these concerns, watermarking techniques have been widely adopted as an efficient means to embed hidden information into digital media, such as images, audio, video, and documents. These watermarks serve as digital fingerprints, enabling content creators to assert ownership, track unauthorized use, and deter potential infringers.

However, with the rise of sophisticated digital manipulation tools, watermark removal and tampering have become prevalent threats, potentially leading to content piracy and unauthorized distribution. Consequently, the demand for robust watermark detection and recognition algorithms has intensified, calling for advanced and reliable methods capable of withstanding various attacks and ensuring the integrity of digital content.

This paper presents a comprehensive review of watermark detection and recognition techniques, highlighting their challenges, advancements, and future perspectives. By examining the evolution of watermarking technology, we aim to provide researchers, practitioners, and policymakers with a comprehensive understanding of the state-of-the-art in watermark detection, setting the stage for future research and innovation in this critical domain. The review covers various aspects of watermark detection and recognition, encompassing both traditional and deep learning-based approaches. We delve into the intricacies of watermark embedding techniques and their impact on detection, exploring the implications of different watermarking strategies on robustness and security. Moreover, we analyze the vulnerability of existing watermarking algorithms to different attacks, including geometric transformations, compression, and signal processing manipulations.

Furthermore, the paper discusses the challenges faced by watermark detection algorithms in real-world scenarios, considering factors like scalability, computational efficiency, and adaptability to diverse media types. We assess the effectiveness of watermark detection methods under various scale and resolution conditions, presenting insights into the limitations and potential improvements in real-time watermark detection. To gain a holistic understanding of watermark detection across different domains, we investigate applications in various industries, such as digital media distribution, multimedia forensics, copyright protection, and intellectual property enforcement. Additionally, we address legal and ethical concerns associated with watermark detection, focusing on privacy implications and the ethical use of watermark detection techniques [1].

This review paper serves as a comprehensive guide to watermark detection and recognition techniques, offering a comprehensive analysis of the challenges faced by existing methods and the strides made in addressing them. As digital media continues to shape our modern world, safeguarding content ownership and intellectual property rights remains paramount. Through a critical assessment of watermark detection techniques, this review sets the stage for future research and innovation, fostering the development of even more robust and reliable methods to protect digital content in an increasingly interconnected world.

## II. CHALLENGES

In the realm of watermark detection, a multitude of challenges must be overcome to ensure the effectiveness and reliability of detection methods. Figure 1 shows the challenges of watermark images. Here, we delve into the significant challenges that watermark detection researchers and practitioners face [2].

- 1. Diverse Nature of Watermark in the Wild:** Watermarks in scene pictures vary and diversify considerably in terms of style, layout, orientation, and other factors. Furthermore, because of the irregular structure, natural scene images contain text from many scripts, complicating the watermark recognition work.
- 2. Complex Imagery Background:** Scene shots typically have very intricate backgrounds, which makes it challenging to identify watermarks in the pictures.
- 3. Interference Factors:** There are various interference factors, including the following:
  - **Uneven lighting:** Uneven lighting is a common issue when photographing in a natural setting due to the varying reactions of sensory devices to the light source. Uneven lighting typically results in incorrect color information and degradation of visual elements, resulting in false detection, segmentation, and identification outcomes.
  - **Blurring and Degradation:** Even in ideal working conditions and with focus-free cameras, watermark blurring and image defocusing can occur. The standard of the watermark is also lowered by additional factors, such as image/video compression and decompression, particularly in the case of graphical video. Character precision is reduced as a result of degradation, blurring, and defocusing, which also causes “touching”, which makes segmentation challenging.
  - **Noise:** Reduced resolution.

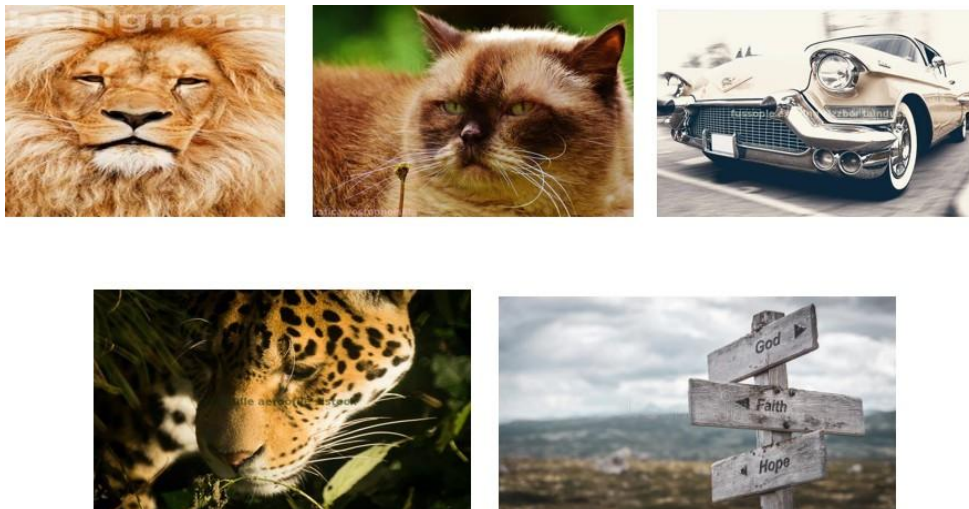
The challenges presented in the realm of watermark detection are formidable but not insurmountable. As technology evolves and adversaries become more sophisticated, it is imperative for researchers, practitioners, and policymakers to collaborate and address these challenges head-on.

## III. APPLICATIONS

To comprehensively understand the significance of watermark detection in the digital age, it's essential to explore its diverse range of applications. From protecting intellectual property to ensuring data integrity, watermark detection plays a pivotal role in various domains, safeguarding authenticity and trust in our digital world [3].

- 1. Copyright Protection:** The legal word for the ownership rights that authors and other artists have over their creative works is "copyright," which is sometimes known as "author's right". The detection of watermark text in images aids in copyright protection by making sure that it cannot be changed or used again without the owner's consent. This suggests that individuals can still inspect a piece before purchasing it without being concerned about a theft.

2. **Fraud and Tamper detection:** Identifying if a certain image is authentic or not is one of the main challenges in image forensics. When images are employed as fundamental evidence to sway judgement, such as, for instance, in a court of law, this might be a significant task. Technically sophisticated digital photography and picture editing software like Adobe Photoshop, Maya, etc. are used to effortlessly modify, manipulate, or tamper with the images without leaving obvious visual indications. In a number of industries, including journalism, scientific journals, digital forensics, medical imaging, and authenticity verification, the abusive use of digital forgeries has become a severe issue. The watermark text detection model will provide a long-term fix to stop producing fake and altered images.
3. **Product Identification:** Product traceability, brand protection, and numerous information labels are all included in the broad area of labelling known as “product identification.” The detection and recognition of watermark text makes it simple to determine whether a product is genuine or original.
4. **Brand Consistency:** The possibility of the assets being exploited is eliminated thanks to watermark identification, which also helps to keep the brand safe. Watermarks can help make it apparent that items are intended to be viewed but not shared.



**Figure 1:** Challenges of watermark images.

#### IV. PREVIOUS SURVEY OVERVIEW

The literature review provides a comprehensive understanding of watermark detection, covering different types of watermarks, various embedding techniques, and a wide range of applications. It forms the basis for analyzing the challenges and advancements in watermark detection and recognition techniques discussed in the paper.

Pal et al.[4] provided the general overview of the various watermarking methods and the media on which they are applied. It examines the difficulties in watermark detection and recognition and establishes the groundwork for watermark detection approaches.

Shi et al. [5] described reversible data hiding techniques are relevant to watermark detection as they involve hiding data without causing any permanent changes to the host media. This survey provides insights into reversible data hiding methods and their potential applications. An overview of the numerous reversible data hiding (RDH) techniques and studies is provided in this publication. The paper divides the RDH schemes into six groups: RDH for images in the spatial domain, RDH for images in the compressed domain, RDH for images that need to be semi-fragilely authenticated, RDH with image contrast enhancement, RDH for images that need to be encrypted, and RDH for images that need to be in both the spatial and compressed domains.

In the Survey of watermarking techniques and applications, Muharemagic et al. [6] explained the digital watermarking, covering the fundamental concepts, various algorithms, and practical applications. It encompasses several watermarking approaches such as spatial domain watermarking, transform domain watermarking, statistical watermarking, and fragile watermarking. Also, it includes discussions on watermark detection and the challenges faced in different scenarios.

## V. WATERMARK DETECTION AND RECOGNITION METHODS

Watermark detection and recognition methods categorized based on common approaches, such as spatial domain, frequency domain, and machine learning-based methods.

Wang et al. [7] proposed a method called GAN to detect and remove potential watermark triggers in the DNN model. There are two phases to it. They initially detect and reverse the watermark in the DNN model using the GAN and a few clean photos. After that adjust the watermarked DNN based on the reversed backdoor photos in the second step. Evaluations carried out experimentally on the MNIST and CIFAR10 datasets.

Deke et al. [8] described a generalized multi-image matting approach that automatically determines the “foreground” (watermark), its alpha matte, and the “background” (original) pictures from a watermarked image collection as input. They investigate and assess the effects of various sorts of irregularities in the embedding of the watermark that may be employed to make watermarking more secure as such an attack depends on the consistency of watermarks across image collections. They used online stock images to illustrate the method, and they offered in-depth quantitative analysis of artificial watermarked data. This paper’s main conclusion is that visible watermarks should be made to be more resilient against mass removal from images rather than only being strong against removal from a single image.

She et al. [9] presented a novel technique for adjusting image similarity and features that, for cross-domain and one-shot watermark detection, beats cutting-edge deep learning techniques. This technique performs well for cross-domain 16,753 class watermark classification performed in a single step. Additionally, based on spatial consistency filtering local matches, they created a matching score and a feature fine-tuning approach. Over strong baselines, this consistency-based method offers significant performance increases.


Independent Component Analysis (ICA) method is used by Yu et al. [10] for watermark detection and extraction process. The original image’s spatial domain is used for watermark embedding. During the Principle Component Analysis (PCA) whitening stage, watermark can be successfully recognized. For blind watermark extraction, a nonlinear robust

batch ICA technique that effectively separates several temporally correlated sources from their observed linear mixes is used. The assessments demonstrate the reliability and efficiency of the suggested ICA-based watermark detection and extraction scheme. The paper has used three types of images: original image, key image, and watermark image. These three images are rearranged into three row vectors to satisfy the input data requirements of the nonlinear blind extraction algorithm-robust batch ICA algorithm, which is used for dewatermarking process. The accuracy of watermark extraction depends on the statistical independence between the original, key, and watermark images and the temporal correlation of these sources.

Jeong et al. [11] proposed a dual watermark detection scheme, which detects watermarks both in spatial and in DCT domains. They used a synchronization method based on block-based template matching to undo the disordered alignment driven on by two significant geometric attacks, rotation and scaling, for the spatial domain detection. In the spatial domain, the watermark’s analogous pattern is extracted. The recovered watermark is then compared against the watermark produced by applying inverse DCT to the original watermark in order to be detected. The suggested dual detection is resilient to both geometric and other signal processing threats, according to the results.

## VI. DATASETS

In this review, we present a comprehensive analysis of datasets commonly used for watermark detection in digital media, ranging from natural images to synthetic water- marked content. Figure 2 shows some sample datasets with image size. As difficulty in collection of data, Shen et al. [9] developed a large public dataset with more than 6k new images. This dataset enables researchers to address at scale two scenarios of practical interest to them: one-shot instance recognition and cross-domain one-shot instance recognition among more than 16k fine-grained classes. They showed that this new dataset is large enough to train contemporary deep learning techniques and that employing mid-level deep features can significantly enhance conventional methods.

Dataset Name	No. of Images	
Adobe Stock	422	
123RF	1376	
CanStock	3000	
fotolia	285	
CVPR17	1000	
Copyrights	1000	
CIFAR10	6000	
MNIST	6000	

Dataset used in [7] and [8]

Example images of Copyright dataset [8]

**Figure 2:** Few dataset samples of watermark detection and recognition.

## VII. CONCLUSION

In conclusion, this comprehensive review has provided a thorough exploration of the watermark detection, recognition, highlighting the evolution of techniques, datasets, challenges, and notable advancements in this crucial field of digital media forensics. We have delved into the various categories of watermarking methods, ranging from robust to fragile, spatial to frequency domain, and investigated their strengths, weaknesses, and applicability in different scenarios.

The analysis of key watermark detection algorithms has revealed the progress made in achieving high accuracy and robustness across diverse watermark types. We've seen how deep learning approaches, such as Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), have revolutionized watermark detection. by leveraging large-scale datasets and complex feature representations. Transfer learning, fine-tuning, and ensemble techniques have further enhanced the effectiveness of these models, enabling them to adapt to new watermarking challenges.

While significant progress has been made, challenges remain. The adversarial nature of watermarking techniques continually pushes researchers to innovate in creating more robust detection methods.

In summary, this review has aimed to provide a comprehensive understanding of watermark detection, from its foundational concepts to state-of-the-art advancements. We hope that this review will be useful to scholars, practitioners, and policymakers navigating the complex realm of digital watermarking and detection.

## REFERENCES

- [1] Safnaz and Bharathi Pilar, "Classification of Watermarked and Non-Watermark Text in Natural Scene Images: An Xception-based Approach", unpublished.
- [2] Potdar, V.M., Han, S., Chang, E.: A survey of digital image watermarking techniques. In: INDIN'05. 2005 3rd IEEE International Conference on Industrial Informatics, 2005., pp. 709–716 (2005). IEEE
- [3] Hussein, E., Belal, M.A.: Digital watermarking techniques, applications and attacks applied to digital media: a survey. *threshold* **5**, 6 (2012)
- [4] Pal, M.: A survey on digital watermarking and its application. *International Journal of Advanced Computer Science and Applications* **7**(1) (2016)
- [5] Shi, Y.-Q., Li, X., Zhang, X., Wu, H.-T., Ma, B.: Reversible data hiding: Advances in the past two decades. *IEEE access* **4**, 3210–3237 (2016)
- [6] Muharemagic, E., Furht, B.: Survey of watermarking techniques and applications. *Multimedia watermarking techniques and applications* **3**(91), 130 (2006)
- [7] Wang, H., Xue, M., Sun, S., Zhang, Y., Wang, J., Liu, W.: Detect and remove watermark in deep neural networks via generative adversarial networks. *arXiv preprint arXiv:2106.08104* (2021)
- [8] Dekel, T., Rubinstein, M., Liu, C., Freeman, W.T.: On the effectiveness of visible watermarks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2146–2154 (2017)
- [9] Shen, X., Pastrolin, I., Bounou, O., Gidaris, S., Smith, M., Poncet, O., Aubry, M.: Large-scale historical watermark recognition: dataset and a new consistency-based approach. In: *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 6810–6817 (2021). IEEE
- [10] Yu, D., Sattar, F., Ma, K.-K.: Watermark detection and extraction using independent component analysis method. *EURASIP Journal on Advances in Signal Processing* **2002**(1), 1–13 (2002)
- [11] Jeong, S., Hong, K., Won, C.S.: Dual detection of a watermark embedded in the dct domain. *EE368A Project Report*