Research Article/Araştırma Makalesi

GAZİ
JOURNAL OF ENGINEERING SCIENCES

Screen Watermark: A Novel Approach in Detecting Digital Criminals

Ömer Faruk Kerman[*a], Aybike Şimşek[b]

Submitted: 18.09.2023 Revised: 22.11.2023 Accepted: 23.12.2023 doi:10.30855/gmbd.0705091

Keywords: Screen Watermarking, Steganography, Data Theft, Malicious Insiders, Investigation and Attribution

[a,*] Kara Kuvvetleri Komutanlığı,
Bilgi Sistem Daire Başkanlığı,
Siber Olaylara Müdahale Şubesi
06100 - Ankara, Türkiye
Orcid: 0009-0005-2871-3867
e mail: ofkerman@kho.msu.edu.tr

[b] Milli Savunma Üniversitesi,
Kara Harp Okulu,
Bilgisayar Mühendisliği Bölümü
06420 – Ankara, Türkiye
Orcid: 0000-0002-1033-1597

*Corresponding author:
ofkerman@kho.msu.edu.tr

ABSTRACT

Organizations need to safeguard their information systems not only against external cyber attackers but also against malicious individuals within their ranks who may exploit their access to steal sensitive information for personal gain. Particularly, these malevolent insiders can covertly monitor and capture documents displayed on computer screens using digital cameras. By opting for digital cameras over more traditional communication methods like email, perpetrators can evade digital traces, making it harder to detect and establish their identity. Even if institutions were to enforce a policy banning the use of camera-equipped devices, the ubiquity of small cameras allows this threat to persist. In response to this menace, a novel technique has been proposed to counter it - embedding hidden watermarks containing confidential information onto the screen, irrespective of the application in use. These invisible watermarks, imperceptible during regular usage, can be extracted from captured images, aiding in pinpointing the location and time of data leaks.

Ekran Filigranı: Dijital Suçluların Tespitinde Yeni Bir Yaklaşım

ÖZ

Kuruluşların bilgi sistemlerini yalnızca harici siber saldırganlara karşı değil, aynı zamanda kişisel çıkar sağlamak amacıyla hassas bilgileri çalmak için erişimlerini istismar edebilecek kendi bünyelerindeki kötü niyetli kişilere karşı da korumaları gerekir. Özellikle içerideki bu kötü niyetli kişiler, dijital kameralar kullanarak bilgisayar ekranlarında görüntülenen belgeleri gizlice izleyebilir ve yakalayabilir. Failler, e-posta gibi daha geleneksel iletişim yöntemleri yerine dijital kameraları tercih ederek dijital izlerden kaçabilir ve böylece kimliklerinin tespit edilmesini ve belirlenmesini zorlaştırabilirler. Kurumlar kamera donanımlı cihazların kullanımını yasaklayan bir politika uygulasa bile, küçük kameraların her yerde bulunması bu tehdidin devam etmesine olanak tanır. Bu tehdide karşı, kullanılan uygulamadan bağımsız olarak ekrana gizli bilgiler içeren gizli filigranlar yerleştiren yeni bir teknik önerilmiştir. Normal kullanım sırasında fark edilemeyen bu görünmez filigranlar, yakalanan görüntülerden çıkarılabilir ve veri sızıntılarının yerini ve zamanını belirlemeye yardımcı olur.

Anahtar Kelimeler: Ekran Filigranı, Steganografi, Veri Hırsızlığı, Kötü Niyetli Çalışanlar, Soruşturma ve İlişkilendirme

To cite this article: Ö. F. Kerman and A. Şimşek, “Screen Watermark: A Novel Approach in Detecting Digital Criminals,” *Gazi Journal of Engineering Sciences,* vol. 9, no. 3, pp. 612-621, 2023. doi:10.30855/gmbd.0705091

1. Introduction

Organizations must safeguard their sensitive information not only from external attackers but also from internal threats, such as infiltrators or malicious insiders. To address this concern, solutions like Data Loss Prevention (DLP) systems are increasingly being employed. However, DLP software can be configured to log or block actions like internet access, sending emails, printing, taking screenshots, or accessing external media. As a result, data leaks through traditional communication channels can either be prevented or at the very least generate a digital trail of the perpetrator's actions. These digital traces can serve as evidence in forensic investigations against malicious insiders. However, solutions like DLP systems cannot prevent an insider from capturing a computer screen's image using a digital camera. Any employee with the authority to view a specific document on the computer screen can capture an image and leak the document's contents to unauthorized parties. Since DLP software cannot detect whether a photo of a document has been taken, using a camera enables the perpetrator to easily evade digital traces. This complicates the process of identifying and proving the perpetrator's identity based on a leaked image. The widespread availability of camera-equipped smartphones and the rise of new technologies like digital glasses or lenses further complicate the control of this data leakage threat.

Techniques such as screen watermarking, steganography, and cryptography offer significant solutions in the field of information security. Screen watermarking involves placing a transparent watermark on the computer screen to facilitate the identification of the source of leaked content through photographs or videos. Steganography is a data hiding technique, referring to concealed data within media files like images, audio, or videos. With this method, detecting hidden data within a document or file becomes challenging. The topic of cryptography encompasses encryption methods and plays a crucial role in data protection. Cryptography prevents unauthorized access by encrypting and transmitting data securely, thereby reducing the risk of data leakage.

2. Information Hiding

Individuals are constantly seeking new and effective ways of communication. Online, users often find themselves needing to send, share, or receive confidential information. As communication increasingly occurs within electronic environments, new needs, challenges, and opportunities emerge. Hence, owing to the rapid development of internet technologies, digital media can be easily transmitted over networks. However, a significant issue encountered when communicating over a network is the presence of multiple observers, including both passive and active ones. While a passive observer might simply listen, an active observer can both listen and modify a message. Consequently, the aim is to ensure that only the intended recipient can decipher the communication's content, while also preferring to keep the transmitted message confidential. To address this issue, two primary solutions have emerged: information hiding and cryptography [1].

The pursuit of a secure and confidential communication method is crucial not only for military purposes but also in terms of market objectives related to commercial strategy and copyright. Cryptography involves the transformation of plain text into encrypted text using a secret key. However, transmitting encrypted text can easily arouse suspicion among attackers, potentially leading to the interception, compromise, or decryption of the encrypted text. To overcome the limitations of cryptographic techniques, the strategy of information hiding has been adopted.

2.1. Information hiding techniques

Information hiding is a multidisciplinary field that enables the concealment of secret data within a digital carrier source. Imagine two parties wishing to communicate while keeping their communication unnoticed by others. The sender can utilize an image that masks the existence of the communication. This image is subsequently made available on a public channel accessible to everyone, yet only the intended recipient is aware of the hidden information and possesses the ability to extract it. Information hiding techniques, as illustrated in Figure 1, are categorized into three main types: steganography, watermarking, and fingerprinting.

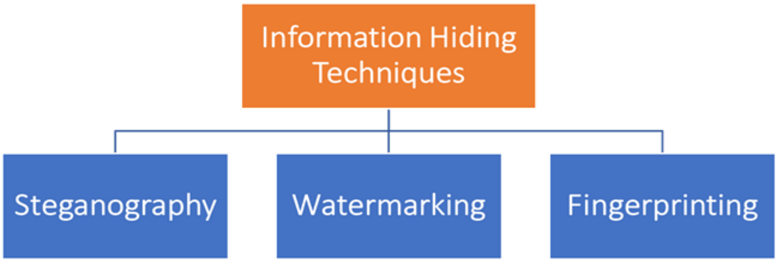

Figure 1. Information Hiding Techniques [1]

Steganography is derived from the Greek words "steganos," meaning "covered or protected," and "graphie," meaning "writing" [2]. As such, steganography is not only the art but also the science of hiding the fact that communication is taking place, along with the actual content of the communication [3]. Privacy is not the sole motivation for steganography. By embedding one data piece within another, they become a single entity, eliminating the need to maintain a connection or risk their separation. An application highlighting this advantage is embedding patient information within medical images, establishing a permanent association between the two pieces of information. The goal of steganography is to enable the transmission of secret messages without arousing suspicion. The concept of "What You See Is What You GET (WYSIWYG)" is sometimes encountered when printing images or other materials, yet it may not always hold true. Images can contain more than what the Human Visual System (HVS) perceives and possess more than a thousand words. Throughout history, people have sought to create covert communication methods.

A steganographic system encompasses two fundamental aspects: steganographic capacity and imperceptibility. However, these two features often conflict and increasing the capacity of a steganographic system while maintaining imperceptibility proves challenging. "Watermarking" refers to identity marks created during the paper-making process. The earliest watermarks emerged in 13th-century Italy and quickly spread throughout Europe, serving to identify skilled papermakers or trade guilds. Today, watermarks still function as markers of origin and for preventing counterfeiting. A watermark is a "hidden message" embedded in a "cover source." Often, extracting a watermark relies solely on knowledge of a secret key [2]. Thus, the effectiveness of any watermarking technique depends on the robustness of the watermark- that is, even if the presence of a watermark is known to exist on a specific object (visible watermarking), it should be impossible to remove the watermark from the object without altering or destroying the original (watermarked) object.

"Fingerprinting" is the process of uniquely marking data to trace the origin of a discovered illegal data copy. The fundamental purpose of fingerprinting is for every user to obtain a copy of the concerned object containing a unique mark. This mark can be used to identify the object and thus the user. For instance, copies could be distributed only to users who authenticate their identity, ensuring they are the ones receiving the copies. Another scenario involves the distribution of sensitive information (images, videos, etc.) to a few authorized individuals and efforts to trace the source of leaked information to a traitorous distributor. These marks should be imperceptible and present in every frame or image distributed. These marks must be embedded so reliably that they cannot be removed through multiple copying or editing processes [2].

2.2. Comparison of steganography and watermarking

Steganography and Watermarking are two distinct concealment technologies utilized for different purposes. Steganography serves the purpose of safeguarding covert communication, whereas Watermarking is employed to authenticate the ownership of documents or content. A comparison table of these two concepts is presented in Table 1.

Table 1. Comparison of Steganography and Watermarking

Features	Steganography	Watermarking
Purpose	To keep communication hidden	To authenticate ownership
Hidden Data	Carried covertly without supervision	Carried consciously or unconsciously
Failure Scenario	Detection of the hidden message	Removal or alteration of the watermark
Output/Result	Stego file containing hidden message	Embedded or invisible Watermark File
Ownership	Does not verify ownership	Verifies ownership
Durability	Plays a minor role	Requires resilience against potential attacks

As shown in Table 1, Steganography focuses on ensuring the undetectability of the hidden message. Hidden data is carried alongside a carrier without supervision, producing an output in the form of a Stego file. Through steganography, it is not possible to determine who the message is from or which organization it belongs to. Additionally, steganography is considered unsuccessful if the hidden message is detected. On the other hand, Watermarking aims to authenticate ownership by adding hidden data known as a watermark to content. The watermark can be visible or invisible and may contain ownership information such as a company logo or owner's details. Watermarking results in a watermarked file, and it is considered a failure if the watermark is removed or altered.

Watermarking is a technology that needs to be resistant against various attacks. This table and explanations showcase the distinct usage scenarios and purposes of Steganography and Watermarking. Steganography is preferred in cases requiring covert communication, while watermarking is chosen when there's a need to authenticate content ownership.

3. Literature Review

This literature review highlights the significance of screen watermarking methods developed through the combination of steganography, watermarking, and cryptography technologies. These techniques are widely employed to protect the confidentiality of sensitive information and authenticate data ownership. Steganographic methods ensure communication privacy by covertly embedding content within other data carriers. Watermarking verifies ownership and provides traceability by adding unique identifiers to content. Cryptography facilitates secure encryption and transmission of data. This study elucidates how these technologies are integrated within the realm of screen watermarking, showcasing their contributions to data security.

3.1. The methods used in information hiding

Data concealment has gained increasing importance in today's world. With the rapid progress of the information age, protecting private and sensitive data has become a significant necessity. In this context, methods like watermarking, encryption (cryptography), and steganography emerge as crucial tools for ensuring data security and guarding against unauthorized access. These methods are highly effective and practical for safeguarding data integrity and security.

Watermarking involves embedding a recognizable image or pattern into documents to establish authenticity. For instance, watermarking is utilized on identity cards, passports, banknotes, and other security documents to prevent counterfeiting and alterations. Digital watermarking, on the other hand, is an embedded marker that aids in protecting digital rights, enabling data to be traced and detected if necessary. Encryption (cryptography) transforms data into an incomprehensible format, offering key features like privacy, integrity, authentication, and non-repudiation. Encrypted data is shielded against unauthorized access and can only be deciphered with the correct key, ensuring data security and preventing unauthorized individuals from accessing the information. Steganography involves the practice of concealing messages, files, or images within other messages, files, or images. This method offers high levels of security and capacity. Steganography hides data without altering the structure of the hidden message, creating a structure that appears like a normal image or file from the outside, but contains concealed information within. This enables data to be transmitted without being detected by unauthorized parties.

These three methods can be combined to achieve a higher level of protection [4]. For instance, messages can first be encrypted into an incomprehensible format. Then, steganography can be employed to embed the encrypted text within a cover medium. This integrated approach successfully meets the goals of security, capacity, and robustness in data concealment. These methods play a crucial role in maintaining data privacy and security. As technology continues to advance, the need for data protection also grows. Therefore, methods like watermarking, encryption, and steganography have become indispensable tools in preserving sensitive data and combating unauthorized access.

3.2. Watermarking methods for screen content protection

Watermarking methods can be categorized into solutions for multimedia files and text documents. Textual watermarks displayed on screens need to be invisible and seamlessly integrated into the text's visuals. Basic approaches for images involve placing watermarks in the least significant bits of pixel values [5]–[7], utilizing imperceptible color changes. However, concerns arise about the ability of devices such as cameras and smartphone cameras to capture these changes. An alternative method [8] encodes watermarks by altering the brightness of adjacent pixels. Unlike this method, our approach doesn't necessitate the presence of the original image for watermark retrieval. Many multimedia watermarking techniques operate in transformed domains, like the frequency spectrum of images [9]–[11], allowing subtle modifications that are hard to detect by humans. However, these watermarks can often be discernible in text documents [12]. Therefore, frequency-domain watermarking isn't well-suited for our proposed method. Similar to our approach, printer technology uses quasi-steganographic techniques, embedding printing details in output using invisible yellow dots [13]. Our method follows a similar logic by hiding information within a watermark on a computer screen.

Considering the above, current research has emphasized three methods. The first approach capitalizes on the human eye's insensitivity to slight brightness changes compared to digital cameras. This content-independent method invisibly embeds watermarks in textual content on screens, aiding forensic investigations into data breaches. The watermark is later extracted from images of the displayed documents, facilitating the determination of breach details. In the watermark embedding process, a unique bit sequence defines hidden data. This sequence is used to create a secure payload with a cryptographic checksum. The payload is encoded using a specialized convolutional encoder to generate watermark symbols embedded into the computer screen. In the extraction process, embedded symbols are decoded to retrieve the protected payload. The cryptographic authentication summary is verified, and if correct, the obfuscated data information is returned; otherwise, no result is provided [14].

The second research paper presents a real-time screen watermarking approach with an overlay layer [15]. It exploits human visual system characteristics to embed an imperceptible watermark over content, adapting to screen changes with minimal delay. This adaptable method functions without specialized hardware and suits various computer systems. The algorithm's simplicity is a key feature, ensuring not only low computational complexity but also enabling real-time processing. The nearly imperceptible delay between screen changes and the watermark algorithm's adaptation is a testament to its high adaptability dynamics. It's worth noting that the delay is proportionate to the size of the watermarked screen, indicating a scalable performance across varying display dimensions.

The third research paper introduces the SSDeN Framework, a frequency-domain watermarking technique leveraging deep neural networks [16]. This framework combats unauthorized use of screen captures without compromising image quality. This innovative framework addresses common limitations observed in traditional watermarking methods, such as detectability and potential degradation of image quality. The SSDeN Framework stands out as a high-performance solution that ensures robust protection against unauthorized screen capture while preserving image quality. It operates in four stages: dataset preprocessing, watermark embedding, separation of watermarked data, and watermark extraction. In the preprocessing stage, data is prepared, compressed, and subjected to Discrete Cosine Transform (DCT). Watermark embedding involves adding the watermark to DCT coefficients corresponding to each pixel. Separation of watermarked data extracts watermarked information using frequency domain attributes. Watermark extraction removes the watermark from watermarked data, recovering the original image. The SSDeN Framework utilizes deep neural networks and frequency domain operations to significantly enhance the security of screen captures without compromising the quality of the captured images. The synergy of these components positions the SSDeN Framework as a cutting-edge solution in the realm of screen capture protection.

3.3. Assessment of watermark methods

In the following examination of various watermarking methods, we delve into the key features, applications, advantages, and limitations of three distinct approaches. The insights from these methods, summarized in the Table 2 below, shed light on the efficacy and challenges associated with contemporary watermarking techniques. In the basis of these methods summarized in Table 2, detailed findings and explanations of the methods used are also presented by us.

Table 2. Comparison of Watermarking Methods for Screen Content Protection

Research Paper	Methodology	Key Features	Application	Advantages	Limitations
[14] Content-Independent Textual Watermarking	Embeds watermarks in textual content on screens, exploiting human eye insensitivity to brightness changes.	Invisibility, aids forensic investigations in data breaches; extraction from images facilitates breach details determination.	Forensic investigations, data breach analysis.	Utilizes human eye characteristics, content-independent.	Effectiveness may be impacted by variations in display technologies.
[15] Real-time Screen Watermarking with Overlay Layer	Real-time screen watermarking using an overlay layer, adapts to screen changes with minimal delay.	Real-time, adaptable without specialized hardware, suits various computer systems.	General screen watermarking applications.	Exploits human visual system characteristics, minimal delay in adapting to screen changes.	Effectiveness dependent on the complexity of screen changes.
[16] SSDeN Framework - Frequency-Domain Watermarking with Deep Neural Networks	Utilizes deep neural networks and frequency-domain operations for watermarking. Operates in four stages: dataset preprocessing, watermark embedding, separation of watermarked data, and watermark extraction.	Combats unauthorized use of screen captures without compromising image quality.	Screen capture security enhancement.	Deep neural network integration, frequency-domain operations, image quality preservation.	Effectiveness subject to real-world scenario tests; potential vulnerabilities not addressed.

The first examined method [14] concluded that all watermark data could be successfully retrieved from unchanged photographs for all tested smartphones. Despite modifications made to the images, the encoded data could still be extracted. Additionally, the method demonstrated remarkable resilience against cropping of watermarked images captured by smartphones. This underscores its efficacy in tracking and analyzing the footprint of protected data, making forensic investigation of data breaches considerably easier.

In the second examined method [15], evaluations were conducted regarding imperceptibility and robustness. When adjusting the size of the watermark region to medium or small, the watermark remained unnoticed in all usage scenarios, with visibility only occurring for larger region sizes. This highlights the balance between visibility level of the watermark and accuracy of placement within regions containing the content to be protected. Robustness assessment revealed that for all test cases with unaltered screen captures, it was possible to remove the watermark. Notably, it was observed that larger region sizes offered greater resistance against attempted attacks, but smaller watermark sizes better withstood processes like cropping.

The third examined method [16] focused on experimental studies. Experiments were conducted on the ImageNet dataset which renowned for its vast collection of labeled images spanning thousands of object categories, serves as a benchmark in computer vision research. In this context, the researchers utilized key metrics such as PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), and MSE (Mean Squared Error) to meticulously assess the performance of the SSDeN Framework. ImageNet, comprising millions of high-resolution images, enables a comprehensive evaluation of the proposed framework's capabilities across diverse visual content, ensuring a robust and thorough analysis of its effectiveness. The outcomes substantiated the effectiveness of the SSDeN Framework as a robust watermarking method. Robustness tests across different scaling and JPEG compression ratios demonstrated the high durability of the SSDeN Framework. This study underscores the potential effectiveness of deep neural network-based watermarking techniques in safeguarding data.

4. Materials and Methods

In the realm of digital security, detecting and preventing data breaches and unauthorized use of screen content have become paramount. While existing watermarking methods address various aspects of this challenge, there is a need for a comprehensive approach that overcomes their limitations. In this section, we propose a novel Screen Watermarking method that not only addresses the shortcomings of the previously discussed methods but also offers an advanced level of security and flexibility in detecting digital criminals. For the implementation of our proposed Screen Watermarking method, we chose the Python programming language. Python offers a rich ecosystem of libraries and tools that facilitate image processing, machine learning, and dynamic behavior integration, all of which are crucial components of our approach. Additionally, Python's readability and ease of use align well with the complexity of the task at hand.

Our chosen methodology integrates content-adaptive watermarking and real-time dynamic overlay integration. This comprehensive approach aims to transcend the limitations of existing methods and provide a robust solution for detecting digital criminals. Content-adaptive watermarking ensures intelligent embedding by considering content characteristics, optimizing visibility while maintaining imperceptibility. The real-time dynamic overlay not only adapts to content changes but also updates based on user interactions, enhancing watermark concealment and confounding digital criminals' attempts at pattern prediction. The workflow of the proposed watermark embedding and extraction process is illustrated in Figure 2.

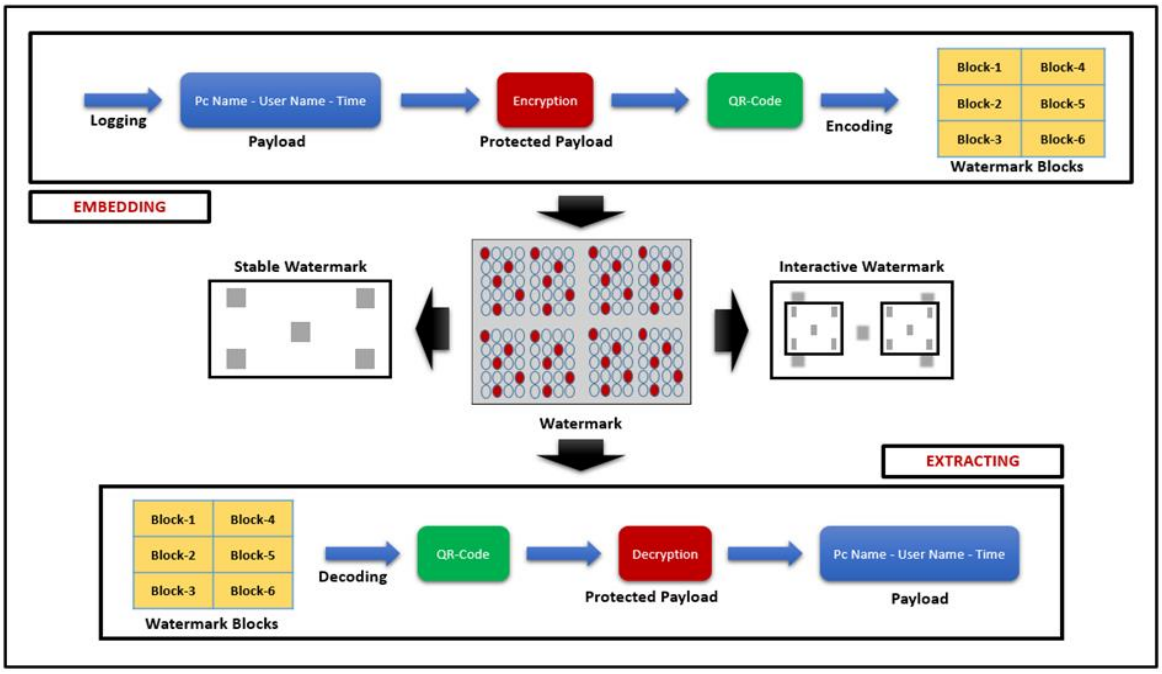

Figure 2. Work Flow of the Proposed Watermarking Process

Through our application, we introduce an innovative methodology for dynamically embedding watermarks into images and real-time inclusion of informative content. Our approach blends image processing, cryptography, steganography, and graphical user interface (GUI) techniques to craft an interactive and visually engaging watermarking solution. Our primary contributions encompass the integration of context-adaptive watermarking and real-time dynamic overlays within a Windows Form application. This enables the generation of QR codes that encapsulate encrypted metadata, subsequently ingrained as watermarks using steganographic techniques in images, complemented by real-time overlays conveying supplementary information. We offer exhaustive insights into our content-adaptive watermarking and real-time overlay methods, supported by a comprehensive step-by-step implementation. This workflow encompasses user input acquisition, encryption leveraging the Advanced Encryption Standard (AES) algorithm, QR code generation via the qrcode library, watermark embedding, and dynamic overlay application. Notably, the content-adaptive watermarking function scrutinizes image features to discern fitting watermark embedding points, while our real-time dynamic overlay function vigilantly tracks the image and applies live overlays. Converting our solution into practice involves employing Python libraries like cv2, numpy, pycryptodome, qrcode, and PyQt5. Our Windows Form application adeptly captures user data, encrypts it, produces QR codes, and subsequently applies both watermarks and real-time overlays. The amalgamation of context-adaptive watermarking with real-time overlays yields captivating and educational images. Through the seamless integration of encryption, QR code generation, and image manipulation within a GUI environment, our proposition showcases its feasibility and potential utility. Forthcoming exploration could delve into intricate watermarking algorithms and diverse overlay contents tailored to varied applications.

5. Discussions

The examination of three different methods raises discussions on the limitations of watermarking techniques in terms of susceptibility to tampering, perceptibility, and robustness, and suggests that their reliability may not be fully guaranteed. These findings indicate that watermark technologies might not offer a complete solution for safeguarding digital content copyright and security.

In the first method [14], it was concluded that watermarks could be retrieved from unchanged photographs. This implies that watermarks can be rendered ineffective. This issue arises as a concern regarding the security insufficiency of watermark technology, potentially failing to fulfill its protective purpose. The study noted that despite alterations to the image, encoded data could still be extracted. This suggests that watermarks can be easily bypassed, allowing unauthorized use of content. This issue could spark a substantial debate on the credibility and efficacy of watermark technology.

In the second method [15], a study was conducted evaluating watermark visibility and robustness. The study revealed that watermarks remained unnoticed when used at specific sizes but were detectable at larger sizes. This underscores the need to strike a balance between watermark visibility and accurate placement within regions containing protected content. However, the study also showed that the watermark could be easily removed from unaltered screen captures. This raises serious concerns about the durability and effective protection capability of watermarking. The limitations and protection deficiencies of this method could be subjects of debate.

The third method [16] examined the SSDeN Framework, a watermarking method. Experimental studies demonstrated the effectiveness of this method. However, as mentioned in the review, the robustness tests were performed under various scaling and JPEG compression ratios. The results concluded that the SSDeN Framework exhibited high durability. Nonetheless, debates might arise about whether these tests fully reflect real-world scenarios. In reality, various attacks and manipulation methods could potentially bypass the watermark. Therefore, further research and discussion may be necessary to ascertain how effective and reliable this method truly is in real-world applications.

The Screen Watermarking approach we propose represents a novel and advanced method for detecting digital criminals engaged in data breaches and unauthorized use of screen content. By addressing the limitations of existing watermarking techniques and integrating innovative features, our approach provides a robust and adaptable solution to the evolving challenges of digital security. The initial prototype, developed as a demonstration within the framework of our proposed methodology, has been executed, revealing successful watermarking of photographs against basic attacks. Based on the initial evaluation results, it was observed that, even after cropping the watermark, 40% of the cropped area retained the capacity to extract the concealed information from the watermarked image. Furthermore, it was noted that the embedded information in the watermark could be recovered even after shooting from various angles and subject to basic photo manipulations, such as brightness adjustments up to 20%. The preliminary assessment indicates that the proposed watermarking method is fundamentally effective. A more comprehensive evaluation will be conducted once the prototype is finalized.

6. Conclusions

This research has culminated in the introduction of an imperceptible low-density watermarking technique for screen content protection. The proposed approach enables the embedding of information that can be subsequently extracted from photographs or screen captures. To further enhance this technique, a coding scheme based on convolutional codes will be developed to accommodate the specific challenges of screen watermarking, including high error rates, non-uniform error distributions, and segmented screen shapes. A series of experiments will underline the resilience of these watermarks against resizing and basic image manipulations, confirming their inconspicuous presence during regular usage.

The forthcoming phase of research will predominantly delve into steganography methods. With the primary objective of creating discreet watermarks on computer screens to counteract data theft without leaving discernible digital traces, the watermarking process will pivot toward steganography techniques. This approach will allow detection information linked to data breaches to be retroactively retrieved from leaked images containing concealed watermarks in the original content. This retrieved information can subsequently serve as compelling evidence in both forensic and administrative inquiries.

Upon the completion of research and the development of watermarking algorithms, a prototype application will be crafted. This application will seamlessly apply invisible watermarks to PC screens and encompass robust mechanisms for detecting data theft. As the final stage, rigorous testing of the developed prototype application will be conducted to evaluate the durability of the watermarking process and the efficacy of data theft detection.

In our future endeavors, we envisage an integrated approach that amalgamates Multi-Layered Watermarking, Behavioral Analysis Integration, and Machine Learning and AI techniques. This unified methodology aims to elevate the watermarking process to unprecedented levels of security and adaptability. By embedding multiple layers of information, introducing behavioral analysis to detect anomalies, and leveraging machine learning to refine watermarking strategies and enhance data theft detection, our proposed approach is poised to stand at the forefront of digital security solutions, ensuring resilient protection against a dynamic landscape of threats.

Conflict of Interest Statement

No conflict of interest was declared by the authors.

References

[1] R. Article, "Available Online at www.ijarcs.info Information Hiding - Steganography & Watermarking : A Comparative Study," vol. 4, no. 4, pp. 165–171, 2013.

[2] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding-a survey," *Proc. IEEE*, vol. 87, no. 7, pp. 1062–1078, Jul. 1999, doi:10.1109/5.771065.

[3] A. D. Orebaugh, "Steganalysis: A Steganography Intrusion Detection System," 2011, [Online]. Available: https://pdfs.semanticscholar.org/0d97/93a6f1a0aef431cebce1b06f37ca4bf99447.pdf

[4] R. Gupta, S. Gupta, and A. Singhal, "Importance and Techniques of Information Hiding : A Review," *Int. J. Comput. Trends Technol.*, vol. 9, no. 5, pp. 260–265, Mar. 2014, doi:10.14445/22312803/IJCTT-V9P149.

[5] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne, "A digital watermark," *Proc. 1st Int. Conf. Image Process.*, vol. 2, pp. 86–90 vol.2, 1994.

[6] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, no. 3.4, pp. 313–336, 1996, doi:10.1147/sj.353.0313.

[7] N. Nikolaidis and I. Pitas, "Robust image watermarking in the spatial domain," *Signal Process.*, vol. 66, pp. 385–403, 1998.

[8] G. Caronni, "Assuring Ownership Rights for Digital Images," 1995.

[9] C. S. Shieh, H. C. Huang, F. H. Wang, and J. S. Pan, "Genetic watermarking based on transform-domain techniques," *Pattern Recognit.*, vol. 37, no. 3, pp. 555–565, 2004, doi:10.1016/j.patcog.2003.07.003.

[10] T. K. M. Tsui, X.-P. Zhang, and D. Androutsos, "Color Image Watermarking Using Multidimensional Fourier Transforms," *IEEE Trans. Inf. Forensics Secur.*, vol. 3, pp. 16–28, 2008.

[11] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia.," *IEEE Trans. image Process. a Publ. IEEE Signal Process. Soc.*, vol. 6, no. 12, pp. 1673–1687, 1997, doi:10.1109/83.650120.

[12] A. M. Alattar and O. M. Alattar, "Watermarking electronic text documents containing justified paragraphs and irregular line spacing," 2004.

[13] M. Embar, L. F. McHugh, and W. R. Wesselman, "Printer watermark obfuscation," in *Proceedings of the 3rd annual conference on Research in information technology*, Oct. 2014, pp. 15–20. doi:10.1145/2656434.2656437.

[14] D. Gugelmann, D. Sommer, V. Lenders, M. Happe, and L. Vanbever, "Screen watermarking for data theft investigation and attribution," *Int. Conf. Cyber Conflict, CYCON*, vol. 2018-May, pp. 391–408, 2018, doi:10.23919/CYCON.2018.8405027.

[15] M. Piec and A. Rauber, “Real-time screen watermarking using overlaying layer,” *Proc. - 9th Int. Conf. Availability, Reliab. Secur. ARES 2014*, pp. 561–570, 2014, doi:10.1109/ARES.2014.83.

[16] R. Bai, L. Li, S. Zhang, J. Lu, and C. C. Chang, “SSDeN: Framework for Screen-Shooting Resilient Watermarking via Deep Networks in the Frequency Domain,” *Appl. Sci.*, vol. 12, no. 19, 2022, doi:10.3390/app12199780.