Journal of Naval Sciences and Engineering
2023, Vol. 19, No. 2, pp. 133-157
Computer Engineering/Bilgisayar Mühendisliği

DOI: 10.56850/jnse.1385482

RESEARCH ARTICLE

**An ethical committee approval and/or legal/special permission has not been required within the scope of this study.*

DEVELOPMENT OF NOVEL COMPARISON BASED STEGANOGRAPHY ALGORITHMS ON MULTIMEDIA TO HIDE PRIVATE DATA

Musa MİLLİ[1*]
Daniyar KHASSENOV[2]

[1]*National Defence University, Turkish Naval Academy, Department of Computer Engineering, Istanbul, Turkiye, mmilli@dho.edu.tr*

[2]*National Defence University, Atatürk Strategic Research Institute, Istanbul, Turkiye, d.khassenov.88@gmail.com*

Received: 03.11.2023 **Accepted: 08.12.2023**

ABSTRACT

Throughout history, humanity has had secrets to safeguard, information that needed to be conveyed to allies in a way that enemies couldn't decipher. To achieve this goal of safeguarding important and valuable information, cryptography and steganography have been frequently used methods both in the past and in today's world. This article introduces a steganographic algorithm designed for hiding data in the color images, along with two different algorithm designs derived from this method. The bits of the hidden message are embedded sequentially into each pixel using the bit comparison. The comparison method works by matching the bits of the image and the message. The least significant bits (LSB) of the carrier (cover) image bytes change depending on the number of matching bits between the carrier image and the hidden message. The proposed method has the potential to hide 1 byte of data within 5 bytes under optimal conditions. The designed algorithms have been tested on a series of color images, and satisfactory results have been achieved in terms of embedding a sufficient amount of data into the images without compromising image quality. The results have been compared with the results of the LSB technique and similar methods based on various performance criteria.

Keywords: *Comparison-based Steganography, Image Steganography, Least Significant Bit, Two-bit Steganography*

MULTİMEDYA ÜZERİNDE ÖZEL VERİLERİ GİZLEMEK İÇİN KARŞILAŞTIRMA TABANLI YENİ STEGANOGRAFİK ALGORİTMALARIN GELİŞTİRİLMESİ

ÖZ

Tarih boyunca insanlığın korunması gereken sırları ve düşmanların çözemeyeceği şekilde müttefiklere iletilmesi gereken bilgileri vardı. Bu önemli ve değerli bilgilerin korunması hedefine ulaşmak için kriptografi ve steganografi hem geçmişte hem de günümüz dünyasında sıklıkla kullanılan yöntemlerden olmuştur. Bu makale, renkli görüntüler üzerinde veri gizleyen stenografik bir algoritma tasarımı ve bu algoritmadan türeyen iki farklı algoritma tasarımını tanımlar. Gizli mesajın bitleri, karşılaştırma yöntemi kullanılarak her bir piksele sırayla gömülür. Karşılaştırma yöntemi, görüntünün ve mesajın bitlerini eşleştirerek çalışır. Taşıyıcı imge ile gizli mesajın eşleşen bitlerinin sayısına bağlı olarak, taşıyıcı imge baytlarının en önemsiz bitleri (LSB) değişir. Önerilen yöntem, optimal koşullar altında 5 bayt içerisinde 1 baytlık veriyi gizleme potansiyeline sahiptir. Tasarlanan algoritmalar bir dizi renkli görüntü üzerinde test edilmiş ve görüntü kalitesinden ödün vermeden yeterli miktarda verinin görüntülere gömülmesi açısından tatmin edici sonuçlar elde edilmiştir. Elde edilen sonuçlar farklı performans kriterlerine göre LSB tekniği ve benzer yöntemlerin sonuçları ile karşılaştırılmıştır.

Anahtar Kelimeler: *Karşılaştırma Tabanlı Steganografi, İmge Steganografisi, En Önemsiz Bit, İki-bit Steganografi*

1. INTRODUCTION

Steganography, derived from the Greek words "steganos" (meaning covered or concealed) and "graphia" (meaning writing or drawing), is the art and science of hiding one piece of information within another. Unlike encryption, which scrambles data to make it unreadable, steganography focuses on rendering data imperceptible to the human eye or automated algorithms. The initial stage on the path to knowledge and wisdom appears as raw data. The advancement of digital technologies has resulted in the creation of more data thanks to numerous IoT devices, scientific studies, social media, e-commerce, and video streams (Li et al., 2022; Milli & Milli, 2023). The generated data has become business intelligence, and as a result, a strategic advantage for those who process it into knowledge and

experience. For these reasons, in today's world, the storage, processing, and protection of data generated in the course of daily life have become increasingly important. Hence, in an age where information is both power and vulnerability, the need for secure communication and data protection has never been greater.

Throughout history, humanity has had secrets to safeguard, information that needed to be conveyed to allies in a way that enemies couldn't decipher. To achieve this goal of safeguarding important and valuable information, cryptography has been a frequently used method both in the past and in today's world. Encryption techniques allow you to secure a message in a manner that prevents outsiders from reading it. Nevertheless, when such a message is transmitted, an external observer will certainly be aware that an encrypted message has been sent, potentially attracting unwanted additional attention for both the sender and the receiver. If the transmitted message appears suspicious, an external observer (attacker) may attempt to decrypt the message, and in some cases, they may succeed. This means that confidential information could be compromised, resulting in its loss of value and significance. In the past, to avoid such situations and conceal the fact of transferring secret information, steganography has often been a frequently employed method. In contrast to cryptography, which aims to conceal the message's contents, steganography aims to hide the message's existence (Bansal & Badal, 2022). In other words, if cryptography makes the understandable unreadable, then steganography makes the visible invisible. The combination of steganography and cryptographic methods is often used, which complements each other.

Computer technologies and digital communication channels are currently advancing in the modern world, and information is frequently presented in the form of multimedia files. Therefore, in addition to older stenography methods, digital steganography has emerged, which uses digital media files to transmit confidential information. Media files are used as containers (carriers) where hidden information is embedded. As a container, it is possible to utilize various types of digital files, such as text documents, audio, images, videos, and so on. Detecting hidden information embedded within digital files is often not easily achievable through

human perception. Therefore, multimedia files containing hidden information can be securely transmitted through public communication channels. Indeed, the real strength of steganography lies in its ability to transmit hidden information through public channels. Because the transmission medium of the hidden multimedia file, the internet, contains a vast amount of similar types of files, the hidden file blends in with them.

The original multimedia file is referred to as the cover object. The embedding function is the technique used to conceal the desired data inside the cover object. To create a stego-object, sensitive data is concealed within the container object using a hiding method. Since the file format of the original object and the stego-object are the same, it is challenging for the human perceptual organs to distinguish any variations. After the stego-object containing the hidden message is transferred to the recipient, the recipient uses an extraction technique to acquire the desired hidden message. Figure 1 shows a general diagram of the steganographic data hiding and extraction data method.

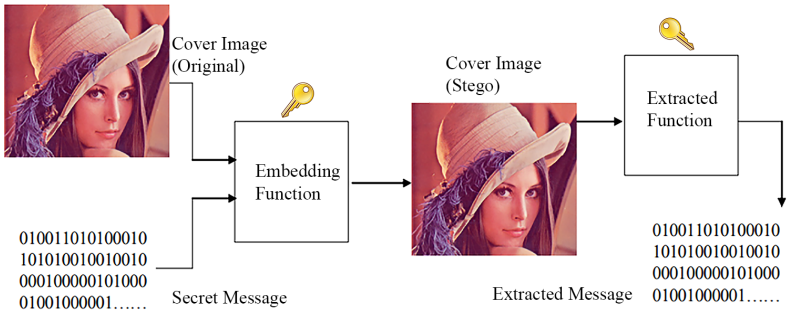

Figure 1. *General diagram of data hiding and data extraction with steganography.*

Some techniques require the original image to access the hidden message from the stego-object. Thus, from the perspective of message retrieval, it is possible to categorize steganography into two distinct classes: those that require a cover object and those that do not. In Figure 2, the classification schema of steganography, adapted from Yalman (Yalman, 2010) can be observed.

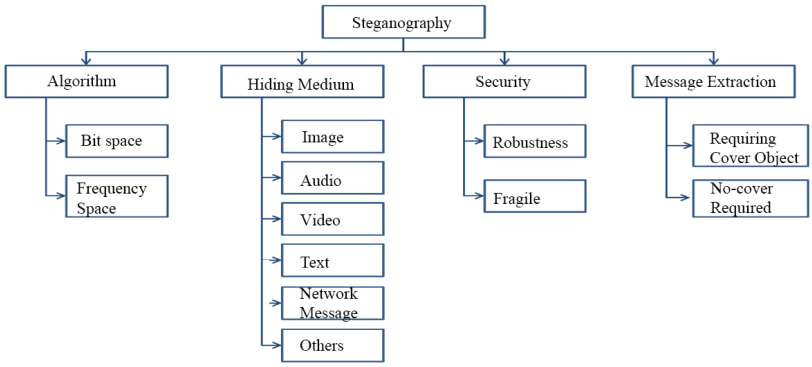

Figure 2. *Classification of stegonography adapted from (Yalman, 2010).*

Three main factors -capacity, robustness, and imperceptibility- determine the effectiveness of steganographic techniques. These factors are known as the "steganography triangle" (Khan & Sarfaraz, 2019). These properties of steganographic objects are often in a trade-off relationship with each other. For instance, when modifications are made to increase the capacity of a particular steganography method, it often negatively affects both robustness and imperceptibility. A steganographic approach is more compatible with classical steganography when its capacity is increased, while it is more compatible with secure steganography when its imperceptibility is increased, and it is more compatible with digital watermarking when its robustness is increased (Grībermans et al., 2016).

Capacity is one of the fundamental characteristics of steganography and refers to how much data a steganographic method can hide. In other words, it indicates the number of bits of secret message that can be concealed. Typically, higher capacity means the ability to hide more secret data. However, achieving a higher capacity may result in a compromise in imperceptibility or robustness, among other important factors. It's crucial to strike a balance between capacity and other features.

Robustness refers to how resistant the hidden message is to operations applied to the carrier medium (e.g., compression or resizing). It measures the ability of the

steganographic method to maintain the integrity of the hidden data even when the cover medium undergoes transformations. A robust steganographic method ensures that the hidden message remains intact despite potential alterations to the carrier. Balancing robustness with capacity and imperceptibility is essential for the overall effectiveness of steganography.

Imperceptibility refers to how well the stego-object (the cover object with hidden data) resembles the original, unaltered cover object. In steganography, it's crucial that any modifications made to the cover object are imperceptible to human perception. If the changes are too noticeable, the steganographic method becomes ineffective because it draws attention to the presence of hidden data. Achieving high imperceptibility while maintaining capacity and robustness is a challenging aspect of steganographic techniques.

The rest of the paper is organized as follows. A thorough analysis of the literature on bit-space and image steganography is provided in section 2. The theoretical designs of the developed methods are explained in section 3. In section 4 findings are presented and some performance criteria are discussed, and finally, the study is wrapped up with conclusions in section 5.

2. RELATED WORK

Image is the most common digital medium used for steganography. Digital images often contain a substantial amount of redundant data, making it easier to conceal a message within the image file, and they also offer ample capacity for this purpose. In digital systems, an image is made up of a numeric series that represents the various light intensities in different regions of the image. The smallest atomic unit of this numerical representation, which is an array, is known as a pixel and these pixels constitute the basis of the image. The steganography of images is made possible by leveraging the limited capabilities of the human visual system (HVS). When a specific digital color in the RGB color format is closely examined, it has been observed that changes in the least significant bits, which have lower significance, are imperceptible to the HVS. Indeed, it is evident that using the

potential of the human eye's perception, a pixel color with a value of (255, 255, 0) cannot be distinguished from a pixel color with a value of (254, 255, 0).

Sharp (Sharp, 2001) is one of the researchers who initially employed the substitution method. According to the approach he developed, unlike the LSB technique, when the last bit of the following byte in the carrier image matches the bit to be hidden, that bit remains unchanged. When the last bit of the next byte in the carrier file is different from the bit to be hidden, the byte value of the container file is either incremented by 1 or decremented by 1. Milikainen (Mielikainen, 2006) is another pioneer who hides data by using the substitution method in his studies. The proposed method operates by hiding 2-bit data in two consecutive pixels in the cover object. Chan (Chan, 2009) improved upon Melikainen's method and proposed a new approach. By hiding one bit per pixel, Chan's proposed technique uses the XOR operator as a function and seeks to affect the image as little as possible. In his proposed method, Tian (Tian, 2003) created a strategy that minimizes image distortion, provides high capacity, and is difficult to reverse. In the proposed method, data is concealed in the area created by doubling the difference between two consecutive pixels in the cover image. Alattar (Alattar, 2004) improved upon Tian's proposed method by enhancing the difference between four consecutive pixels by a factor of two, allowing for the concealment of 3-bit data in the resulting area. Chang et al. (Chang et al., 2009) created two distinct stego-images in their developed approach. In the developed method, data is concealed in two images created using a modulation matrix and a change direction. The method exhibits a substantial data concealment capacity due to its utilization of two images. Lu et al. (Lu et al., 2015) proposed an alternative method by enhancing the approach suggested by Chang (Chang et al., 2009). The method provides high data hiding capacity alongside high image quality in pixel container images. In the study conducted by Ker (Ker, 2004), three pixels are used to conceal two data bits. While the last bits of two pixels are used for data hiding, the third one indicates whether the embedded image contains hidden data or not. Wu and Tsai (Wu & Tsai, 2003) proposed a method of concealing data using the pixel differencing technique. In this method, differences in pixel values are calculated in

a way that consecutive pixels in the container image do not overlap. The possible values of differences are represented by different classes. Building on the research of Wu and Tsai (Wu & Tsai, 2003), Wang et al. (C.-M. Wang et al., 2008) offered a novel approach. The method operates based on concealing data using the result of the modulus function of the difference between two pixels. Experimental results have revealed that the algorithm developed by Wang et al. outperforms previously suggested algorithms in the literature (Saran & Olcay, 2013).

In addition to the traditional methods described above, there are also state-of-the-art studies available. Two distinct steganographic approaches are described in Swain(Swain, 2018). One of these approaches works by dividing the image into 2x3 pixel blocks, while the other one operates with 3x3 pixel blocks. Only one pixel's value is altered within each block, and the difference values of neighboring pixels are calculated with respect to the changed pixel's new value. The method is more challenging to detect using steganalysis techniques compared to traditional methods because it makes decisions based on the pixel values of the cover image, like our proposed method.

In the study conducted in 2021, Durdu (Durdu, 2021) carried out research on image hiding within images, distinct from the studies mentioned above. In the proposed study, 24-bit color images are divided into 4-bit chunks, and each 4-bit chunk is reduced to 2 bits. In the process of retrieving the hidden image, the 2-bit data is reversibly transformed into 4 bits, and the 24-bit image is reconstructed by combining them in this manner. Given that 2-bit representations of 4-bit image blocks exist, some data loss is inevitable. However, due to the reduction in the size of the data to be concealed, twice as much data can be hidden compared to the traditional LSB method.

The structure of the cover object has a significant impact on how well data can be hidden within a digital item. Due to this, several researchers (Volkhonskiy et al., 2020; Zi et al., 2018) have created images using artificial intelligence methods that are better suited for data concealment as opposed to using already-existing images. In their 2020 study, Volkhonskiy et al. (Volkhonskiy et al., 2020) generated cover

images using the Generative Adversarial Networks (GANs) technique and concealed data into the generated image using the LSB embedding algorithm.

3. THE PROPOSED SUBSTITUTION BASED METHODS

Three different data-hiding techniques based on the comparison approach have been developed in the paper, and bitmap (BMP) images were used for the experimental evaluation of these techniques. The three developed methods have been named sequentially as comparison method 1 (CM-1), comparison method 2 (CM-2), and comparison method 3 (CM-3).

3.1. 2-bit comparison-based steganography method (CM-1)

RGB-based image files consist of pixels, with each pixel having three different values, each comprised of 8 bits. So, a pixel contains 24 bits of color data, organized as 8 bits for Red, 8 bits for Green, and 8 bits for Blue, in that order. In the classic LSB method, a single bit of data is embedded into the least significant bit (LSB) of each 8-bit color value, representing a pixel. Thus, each pixel conceals 3 bits of data.

The subject byte of the cover image is divided into two fields called the matching field and the position field. The matching field has six bits consisting of three-bit pairs, and the position field has two bits. The developed 2-bit comparison method (CM-1) sequentially compares 2 bits of hidden information with the bit pairs of the matching field and the match information is stored position field if the bit pairs to be hidden match any two bits in the cover image. As a result, the created solution only slightly modifies the cover image.

As seen in Table 1, the CM-1 method divides each hidden information byte into four to create matched bits. Additionally, by taking the first pair of a byte of hidden information, it compares it in pairs with the image bits. If the first two bits of the cover image byte match the two bits to be hidden, the last two bits of the cover image are set to 00; if it matches with the second pair, they are set to 01, and if it matches with the third pair, they are set to 10. If there is no match, then the last two bits are set to 11, and the value of the unmatched pair of hidden information bits

replaces the 5th and 6th bits. b_i and b_{i+1} are the first and second bits of hidden data, respectively. This way, to be used in the process of extracting hidden data, the position of the stored information in the stego-image is indicated in the last two bits. The working principle of the CM-1 method is illustrated in Figure 3.

Table 1. *The identity table and position values of the CM-1 method.*

The comparison table between the bit pairs of the hidden message and the bit pairs of the cover image.	The configuration of the final two bit pairs in every byte of the stego-image. (position field)
first pair (1st and 2nd Bit)	0 0
second pair (3rd and 4th Bit)	0 1
third pair (5th and 6th Bit)	1 0
no match	b_i b_{i+1} 1 1 (The unmatched bits of the hidden message are written in place of the 5th and 6th bits of the cover image.)

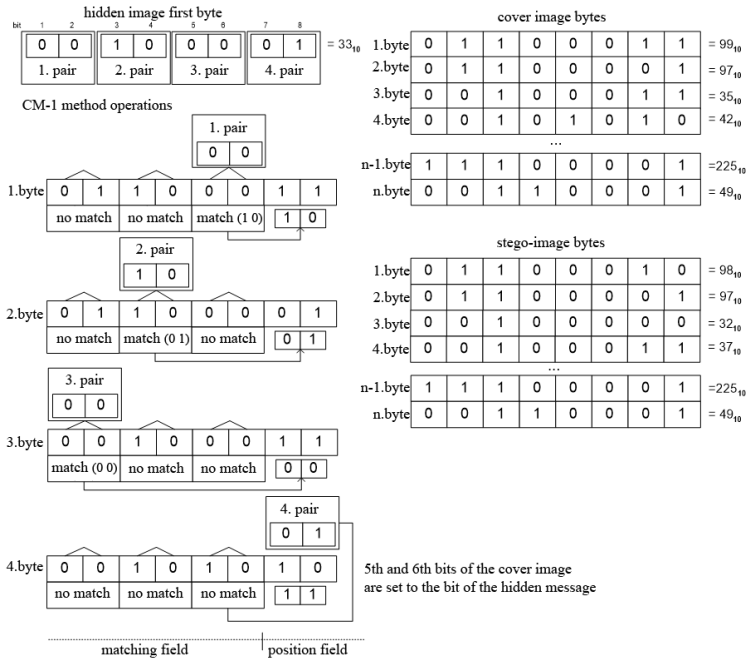

Figure 3. *Operating concept of the CM-1 method.*

Although the location for concealing data may work stochastically with respect to the structure of the cover image, the capacity is determined deterministically and remains constant for any cover image. In other words, the CM-1 method can embed 1 byte of data into 4 bytes of cover-image. However, the rate of change in the cover image varies stochastically depending on the cover image structure itself. The cover image will not change in the best-case scenario if the concealed message's bits match the matching field bits and the last two bits remain unchanged. Otherwise, the 5th and 6th bits of the image byte change resulting in maximum alterations in the stego-image. With the CM-1 method, an alteration ranging from 0% to 6.25% can occur in the image.

3.2. 4-bit comparison-based steganography method (CM-2)

In the CM-2 method, similar to the CM-1 method, the procedure of comparing similar bits and specifying their positions is followed. However, in this method, the bytes of the hidden message are not divided into pairs of 2 bits. Instead, in the CM-2 method, the hidden data is divided into 4-bit chunks, and comparisons are made with 4-bit segments of the cover image. Unlike the CM-1 method, even though 4-bit comparisons are made, it is not expected that all 4 bits will match. The matching process stops at the bit where the match ends, and it does not continue. That is, starting with the first bit of the secret message, it is compared to the cover image bits sequentially. If the bits match, one bit is added at a time and 4 bits in the chunk are tried to be completed. In the worst case, when the bits of the secret message do not match the bits of the cover image, the still valid byte of the cover image is used to conceal the bits of the message. Table 2 illustrates the bit conditions of the cover image, both in situations where matching takes place and when it does not. Here, the b_i value represents the next bit to be concealed.

When the proposed CM-2 algorithm is compared to the CM-1 algorithm, it can be observed that the maximum data-hiding potential of the CM-2 algorithm is lower. Although the CM-2 algorithm offers a lower capacity, at 30%, its detectability has significantly decreased compared to the CM-1 algorithm and traditional algorithms.

In other words, the CM-2 method results in such a minimal visual difference between the stego-image and the original image that it is nearly imperceptible.

Table 2. *The identity table and position values of the CM-2 method.*

The comparison table between the bit pairs of the hidden message and the bit pairs of the cover image.	The configuration of the last three bits in every byte of the stego-image. (position field)
1st bit	0 0 1
2nd bit	0 1 1
3rd bit	1 0 1
4th bit	1 1 1
no match	b_i 0

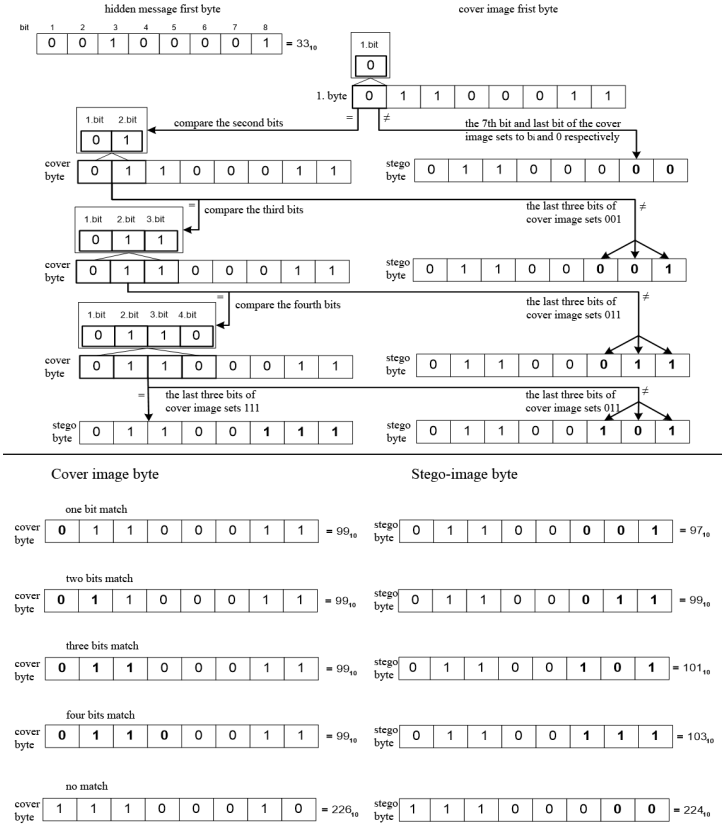

Figure 4. *Operating concept of the CM-2 method.*

In Figure 4, the working principle of the CM-2 method is illustrated. In the process of hiding data, the first step is to check for a match between the leftmost bit of the hidden message and the first bit of the cover image. If the first bits are not equal, the 7^{th} bit of the cover image byte is set to the hidden message bit, and the last bit is set to 0 (zero). In cases of one-bit, two-bit, three-bit, and four-bit matches, the last three bits of the cover image are set to 001, 011, 101, and 111, respectively.

The proposed CM-2 method can potentially cause a maximum of 3-bit changes for 4-bit hidden data. Besides, the CM-2 method has the potential to hide 1 byte of data within 2 bytes of data. Although the traditional LSB method changes 1 bit for 1 bit of data, it can embed 1 byte of data into 8 bytes of the cover image. Taking into account the bit values, while the CM-2 method results in 7 times more changes compared to the LSB method, it can embed 4 times more data than the LSB method in the ideal scenario.

3.3. 3-bit comparison-based steganography method (CM-3)

The CM-3 method follows a similar approach to the procedures described above, which involve comparing similar bits and specifying their positions. The cover image, however, is separated into 2-byte chunks in this technique. The first three bits of the secret message byte are matched with the first three bits of the first byte of the cover image, respectively, and it is checked whether they are equal. Information about the number of matched bits or non-matching situations is written to the 8^{th} bit of the 2-byte chunk in question. Table 3 illustrates how the cover image bytes change in matching and non-matching situations, and Figure 5 provides a schematic representation of the CM-3 method's operation.

In comparison to the previously described CM-1 and CM-2 algorithms, the CM-3 technique conceals a lower quantity of data. Experimental studies have shown that the capacity provided by the CM-3 algorithm is 64% lower than that provided by the CM-1 algorithm. In contrast, the criteria of imperceptibility, one of the quality criteria for steganographic objects, has increased. Furthermore, the CM-3 method is competitive in terms of all three quality criteria (capacity, robustness, and imperceptibility) when compared to previous studies.

Table 3. *The identity table and position values of the CM-3 method.*

Comparison between the bits of the hidden message and the first byte bits of the carrier image in each chunk.	The configuration of the last bits in every byte of the stego-image chunks. (position field)
1st bit	last bit of 1. byte = 0 last bit of 2. byte = 0
2nd bit	last bit of 1. byte = 0 last bit of 2. byte =1
3rd bit	last bit of 1. byte = 1 last bit of 2. byte = 0
no match	last bit of 1. byte = 0 last two bits of 2. byte = b_i 1

Figure 5 explains the flowchart of the CM-3 method. In the data hiding process, the first step involves checking the equality of the first bits of the hidden message and the cover image's first bytes. In case of no match, the last bits of the cover image's first and second byte are set to 1, and the 7th bit of the second byte is set to the non-matching bit of the secret message. When the first bits match, the matching of the second and third bits is checked in order, and the last bits of each of the two bytes in the cover image are set to 00, 01, and 10 respectively.

The CM-3 method accomplishes 2-bit modifications in the cover image for hidden messages ranging from 1 to 3 bits. In other words, under ideal conditions, the suggested CM-3 method allows for the hiding of 1 byte of data inside 6 bytes of cover image. While theoretically, the CM-3 method can store up to 32.76 KBytes of text within a 256x256 pixel image, experiments have shown that on average, around 16 KBytes of data can be stored in a random image. The traditional LSB method allows for the storage of up to 24.57 KBytes of data within the same size image. Considering the performance criteria of steganographic methods, it becomes evident that the developed CM-3 method holds the potential to compete with traditional LSB methods and other counterparts.

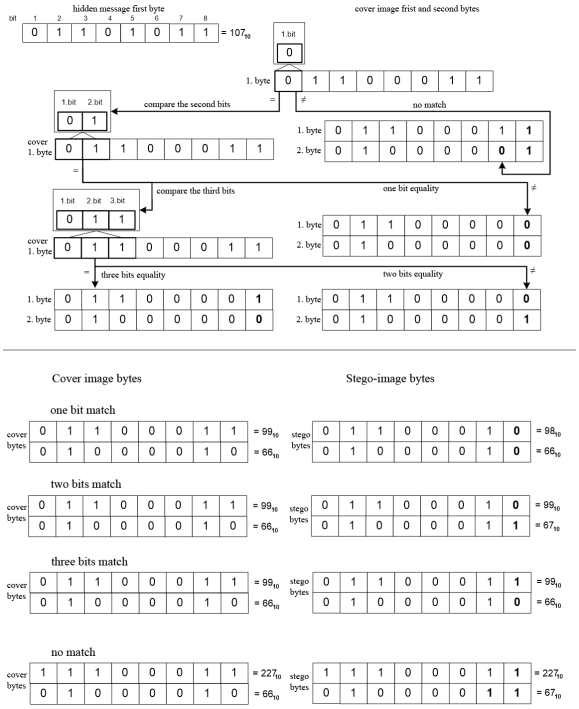

Figure 5. *Operating concept of the CM-3 method.*

4. EXPERIMENTAL RESULTS AND DISCUSSION

This section discusses the performance of the proposed approaches using experimental findings. The proposed methods were tested on a variety of images with different characteristics (standard, complex, plain-colored, etc.) commonly used in the field of steganography. Several outcomes have been provided in order to illustrate the efficacy and performance of the developed methods as measured by commonly used metrics: Peak Signal-to-Noise Ratio (PSNR) and Mean Squared Error (MSE).

The maximum number of bits that may be embedded in a fixed-sized image is commonly characterized as image capacity. Proposed methods, hide the data considering the bits of the cover image, which makes it challenging to determine

capacity in a general sense, as it depends on the bit structure of the cover image. However, theoretically, one can discuss the maximum and minimum capacity for the proposed methods. Given that capacity is one of the key performance criteria in steganography algorithms, it has been considered when discussing the findings.

In many studies in the literature, criteria that involve comparing the stego-image with the original image are frequently used to detect any distortions in steganographic images after the data-hiding process. MSE (Equation 1), PSNR (Equation 2), and Structural Similarity Index Measure (SSIM) (Equation 3) (Z. Wang et al., 2004) are commonly used evaluation measures to assess the quality and extent of changes in steganographic images. SSIM is a numerical method for calculating the similarity of two images. The SSIM index is a full-reference metric, meaning that the measurement or estimation of image quality relies on a reference, typically an uncompressed or undistorted original image. In Equation 1, *m* and *n* represent the row and column information of the image; *O* represents the original image, and *S* represents the stego-image. The PSNR is calculated once the MSE value has been obtained. *MAX* denotes the maximum value a pixel may have in Equation 2, which is 255 for RGB. In Equation 3, SSIM is calculated for different image windows. Where μ_x and μ_y are the pixel values mean of *x* and *y* images respectively. The μ_x^2 and μ_y^2 are the variance values of *x* and *y* images respectively. The σ_{xy} is the covariance of *x* and *y*. c_1 and c_2 are the variables to stabilize the division with weak denominator.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[O(i,j) - S(i,j)] \quad (1)$$

$$PSNR = 10log_{10}(\frac{MAX^2}{MSE}) \quad (2)$$

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (3)$$

The Lena (640x480), Bird (500x400), Baboon (400x300), and Peppers (640x480) images, which have been commonly used in previous research studies, were also selected for experimental purposes in this study. The stego-images obtained by hiding 500 bytes of data using the developed methods and the original images are

depicted in Figure 6. Examining the created stego-images and the original images, no changes can be perceived by human sensory organs.

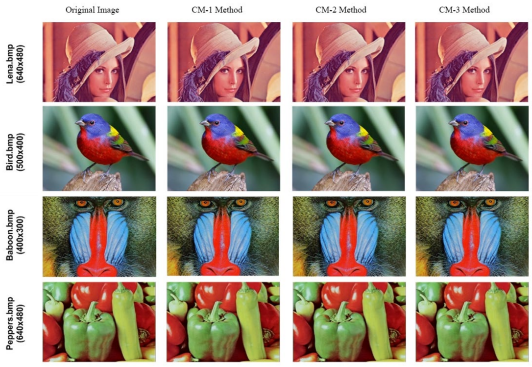

Figure 6. *Stego-images with 500 bytes embedded and the original images.*

In Table 4, a comparison of the results of the proposed methods and the work of Wu and Tsai has been conducted using the MSE and PSNR metrics. Developed CM-1, CM-2, and CM-3 methods, which can embed a substantial amount of hidden data, have yielded MSE and PSNR values comparable to their counterparts. The PSNR values obtained for the proposed methods demonstrate that the degradation in the stego-image is minimal. The CM-3 method, in particular, is more effective in terms of PSNR values despite having a lesser data embedding capacity than the CM-1 and CM-2 methods.

A high PSNR value obtained after data hiding implies a high degree of similarity between the cover image and the original image. In other words, a higher PSNR value indicates that the steganography method has made fewer changes to the cover image. Small MSE values, on the other hand, suggest that the method is better. All three developed methods embed data depending on the structure of the cover image, as seen in Table 4, substantial PSNR variations are observed for different images. Upon evaluating the findings, it is evident that the CM-3 method demonstrates superiority over the other methods (CM-1, CM-2, and Wu & Tsai). Furthermore, MSE values have been calculated for the stego-images generated after the hiding process for each compared method. The results obtained by hiding

messages of different sizes in four different images used in experimental studies are shown in Table 4, and the average of these results is presented in the following figures. The figures depicting the average MSE and PSNR values calculated for each method are provided in Figure 7 and Figure 8, respectively. Considering the averages, the results also establish that the CM-3 method offers the best outcome.

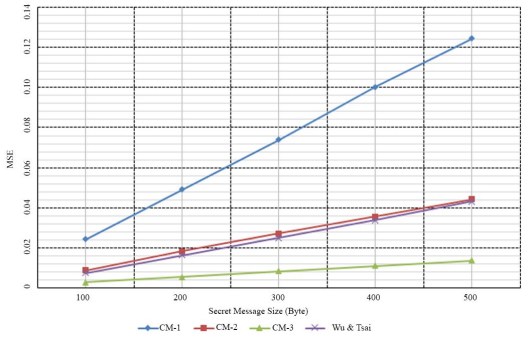

Figure 7. *The average MSE values of stego-images produced by hiding secret data of varying sizes.*

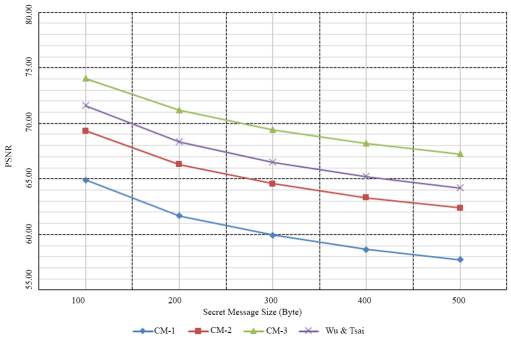

Figure 8. *The average PSNR values of stego-images produced by hiding secret data of varying sizes.*

In the parametric analysis, one of the measured values is the changing number of bits. Table 5 presents the number of bits changing obtained from the parametric analysis. Based on this experimental investigation, it is evident that the stego-image obtained using the CM-1 approach has the least amount of bit modification.

When the bit alteration parameter was considered, the proposed techniques CM-2 and CM-3 produced inferior results than the CM-1 method. However, they have produced better results compared to the method proposed by Wu & Tsai.

Table 4. *Comparison of proposed and equivalent methods with each other.*

	Lena.bmp		Bird.bmp		Baboon.bmp		Peppers.bmp		Method
	MSE	PSNR(db)	MSE	PSNR(db)	MSE	PSNR(db)	MSE	PSNR(db)	
100	0.0106	67.8445	0.0192	65.2823	0.0263	63.9266	0.0108	67.7897	**CM-1**
	0.0035	72.5937	0.0060	70.2929	0.0104	67.9327	0.0038	72.2446	**CM-2**
	0.0014	**76.5344**	**0.0017**	**75.6266**	**0.0033**	**72.8875**	0.0019	75.1970	**CM-3**
	0.0014	76.4026	0.0026	73.9120	0.0109	67.7543	**0.0018**	**75.3682**	**Wu & Tsai**
200	0.0216	64.7689	0.0344	62.7618	0.0525	60.9285	0.0336	62.8570	**CM-1**
	0.0071	69.6038	0.0129	67.0069	0.0207	64.9687	0.0071	69.6171	**CM-2**
	0.0029	**73.4305**	**0.0035**	**72.6818**	**0.0066**	**69.8989**	**0.0031**	**73.1567**	**CM-3**
	0.0033	72.9190	0.0053	70.8149	0.0282	63.6278	0.0037	72.3554	**Wu & Tsai**
300	0.0312	63.1768	0.0504	61.0986	0.0823	58.9741	0.0506	61.0809	**CM-1**
	0.0108	67.7818	0.0187	65.3923	0.0318	63.0974	0.0106	67.8489	**CM-2**
	0.0042	**71.8655**	**0.0053**	**70.8744**	**0.0099**	**68.1513**	**0.0049**	**71.2238**	**CM-3**
	0.0048	71.2431	0.0082	68.9873	0.0454	61.5570	0.0058	70.4764	**Wu & Tsai**
400	0.0415	61.9416	0.0726	59.5195	0.1100	57.7164	0.0632	60.1225	**CM-1**
	0.0146	66.4799	0.0250	64.1467	0.0417	61.9199	0.0153	66.2738	**CM-2**
	0.0055	**70.6928**	**0.0071**	**69.6039**	**0.0133**	**66.8787**	**0.0065**	**69.9386**	**CM-3**
	0.0066	69.9030	0.0110	67.6919	0.0636	60.0909	0.0073	69.4390	**Wu & Tsai**
500	0.0515	61.0122	0.0901	58.5803	0.1349	56.8287	0.0720	59.5553	**CM-1**
	0.0181	65.5410	0.0307	63.2525	0.0521	60.9566	0.0185	65.4559	**CM-2**
	0.0068	**69.7539**	**0.0088**	**68.6434**	**0.0167**	**65.8813**	**0.0079**	**69.1186**	**CM-3**
	0.0083	68.9339	0.0140	66.6581	0.0822	58.9776	0.0092	68.4550	**Wu & Tsai**

Table 5. *The changing numbers of bits obtained through the parametric analysis of the compared methods.*

	Lena	Bird	Baboon	Peppers
	640 * 480	500 * 400	400 * 300	640 * 480
CM-1	**3037**	**3096**	**3092**	**3333**
CM-2	3592	3585	3652	3700
CM-3	3772	3508	3654	3965
Wu & Tsai	3999	4321	3777	4101

One of the essential assessment criteria for a steganographic method is the amount of data it can hide. As a result, the proposed algorithms and traditional methods have been assessed based on their data-hiding potential using the images employed in experimental studies. The findings are presented in Table 6, and based on the results, the CM-1 method demonstrates the highest capacity. The developed CM-1 method has the potential to embed two times as much data compared to the classic LSB method and three times as much data compared to Wu & Tsai's method. There is no doubt that the discussed criteria are theoretical potentials, but appropriate images for the method can be found or generated using artificial intelligence tools.

Table 6. *Comparison of the maximum bit embedding capacities of the developed methods and traditional methods.*

	Lena	**Bird**	**Baboon**	**Peppers**
Size	640 * 480	500 * 400	400 * 300	640 * 480
LSB	116 350 byte	75 750 byte	45 450 byte	116 318 byte
CM-1	**232 704 byte**	**151 500 byte**	**90 902 byte**	**232 718 byte**
CM-2	163 840 byte	106 166 byte	62 208 byte	158 671 byte
CM-3	77 467 byte	51 106 byte	30 261 byte	76 962 byte
Wu & Tsai	67 129 byte	44 638 byte	29 951 byte	66 598 byte

In addition to the advantages of the developed algorithms such as high capacity and imperceptibility, their detection by steganalysis methods is equally challenging. Traditional LSB and LSB-based methods embed one or two bits per byte in the cover image using a deterministic approach, whereas the three developed methods embed hidden messages in locations that vary according to the structure and bit sequence of the cover image. Thus, it is impossible to predict where the developed method will embed any bit of the hidden message without knowledge of the algorithm. Indeed, the regions in which the CM-2 and CM-3 methods made changes on a bird image can be clearly seen in Figure 9. The images at the top in Figure 9 belong to the CM-2 method. The first one represents the original image, the second one the stego-image, and the subsequent images illustrate where the bit corruptions are prevalent. Similarly, in Figure 9, the images at the bottom belong to the CM-3 method. Examining areas with intense bit corruption, it becomes evident

that predicting where the corruption will occur is not feasible. From this perspective, the developed methods outperform traditional LSB-based techniques.

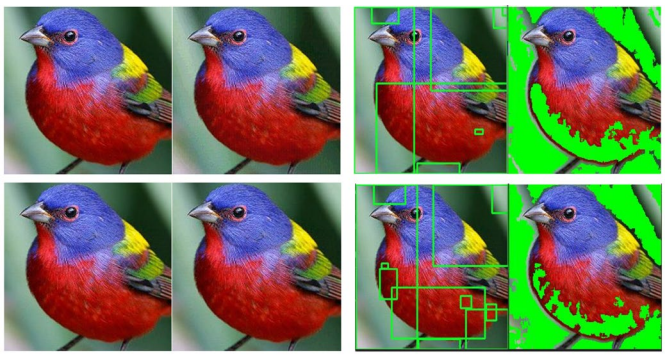

Figure 9. *Original images, stego-images, and regions with corruptions (the images above belong to CM-2, the images below belong to CM-3).*

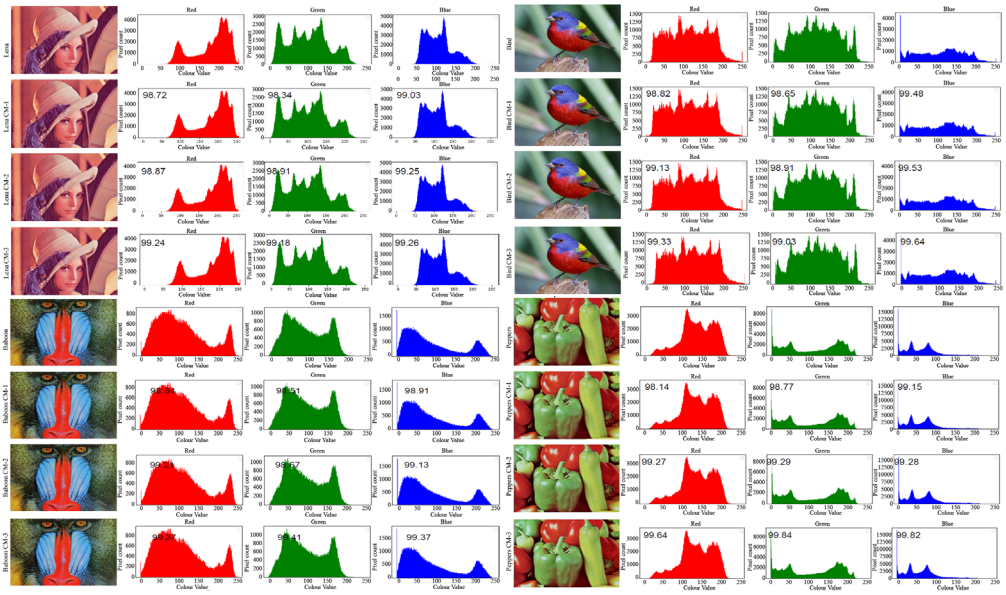

Figure 10. *Comparison of histograms between the original and stego-images.*

The frequency graph of pixel intensities in an image is referred to as a histogram. A histogram provides information about how many pixels have each color value in an

image. During data embedding, the pixel value changes, causing a change in the color value as well. It's a measure used in steganography and indicates how much the color values of the stego-image have changed. It is expected that the color values remain unchanged or that any changes are minimal. Figure 10 displays the results of an experimental study that illustrates how much the color histograms have changed after the process of embedding 500 bytes of data into different images using the three developed methods. In this experiment, the CM-3 method has achieved results that are closer to the original image's color values for each image. This indicates that the CM-3 method made fewer changes in the image when considering the color histogram.

5. CONCLUSION

There are many different steganography methods that work with images in various formats, both color and grayscale. Many approaches frequently employ sophisticated mathematics to produce effective outcomes. Some of these methods use the original image to extract hidden data. The study has proposed new and effective methods for hiding data in images without causing noticeable alterations. In the proposed method CM-1, data up to 3 bits can cause changes within 2 bytes. In other words, the CM-1 has the potential to hide 1 byte of data within 5 bytes under optimal conditions. The proposed methods use similarity and comparison tactics without complex mathematical calculations. This makes them time-efficient, as they don't involve high-time complexities. Moreover, the proposed methods are relatively simple and straightforward to apply to RGB color images compared to some of their counterparts. Even if the stego-image is obtained by third parties, it is not possible to extract the hidden data without knowledge of the algorithm. To extract embedded data from a stego-image, there is no need to refer to the original image.

Based on the results obtained from the experimental studies, it is evident that the developed CM-3 method achieved more effective results in terms of the PSNR and MSE metrics. There is a difference of at least 4.5 units between the results of the CM-3 method and the Wu and Tsai algorithm. However, in terms of data

embedding capacity, the leading method in this field with the best results is CM-1. In steganography, there is a trade-off between capacity and its imperceptibility. When this trade-off is examined in the context of the developed methods, it is evident that CM-1 has a high data-hiding capacity while having relatively lower imperceptibility. Conversely, CM-3 exhibits the opposite characteristics, with a low data-hiding capacity and high imperceptibility.

The experimental findings reveal that the suggested approaches preserve the acceptable visual quality of stego-images while adequately fulfilling the requirements of steganography.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

REFERENCES

Alattar, A. M. (2004). "Reversible watermark using the difference expansion of a generalized integer transform." *IEEE Transactions on Image Processing*, *13*(8), 1147–1156.

Bansal, R., & Badal, N. (2022). "A novel approach for dual layer security of message using Steganography and Cryptography." *Multimedia Tools and Applications*, *81*(15), 20669–20684. https://doi.org/10.1007/s11042-022-12084-y

Chan, C.-S. (2009). "On using LSB matching function for data hiding in pixels". *Fundamenta Informaticae*, *96*(1–2), 49–59.

Chang, C.-C., Chou, Y.-C., & Kieu, T. D. (2009). Information hiding in dual images with reversibility. *2009 Third International Conference on Multimedia and Ubiquitous Engineering*, 145–152. https://ieeexplore.ieee.org/abstract/document/5319030/

Durdu, A. (2021). "24-bit renkli imge içine 24-bit renkli imge gizleyen yüksek kapasiteli düşük bozulumlu tersinir kayıplı yeni bir veri gizleme yöntemi (YKKG)." *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, *27*(2), 96–113.

Grībermans, D., Jeršovs, A., & Rusakovs, P. (2016). "Development of Requirements Specification for Steganographic Systems." *Applied Computer Systems*, *20*(1), 40–48. https://doi.org/10.1515/acss-2016-0014

Ker, A. D. (2004). Quantitative evaluation of pairs and RS steganalysis. *Security, Steganography, and Watermarking of Multimedia Contents VI*, *5306*, 83–97. https://www.spiedigitallibrary.org/conferenceproceedingsofspie/5306/0000/Quantitativeevaluation-of-pairs-and-RS-steganalysis/10.1117/12.526720.short

Khan, A., & Sarfaraz, A. (2019). "Novel high-capacity robust and imperceptible image steganography scheme using multi-flipped permutations and frequency entropy matching method." *Soft Computing*, *23*(17), 8045–8056. https://doi.org/10.1007/s00500-018-3441-1

Li, C., Chen, Y., & Shang, Y. (2022). "A review of industrial big data for decision making in intelligent manufacturing." *Engineering Science and Technology, an International Journal*, *29*, 101021. https://doi.org/10.1016/j.jestch.2021.06.001

Lu, T.-C., Tseng, C.-Y., & Wu, J.-H. (2015). "Dual imaging-based reversible hiding technique using LSB matching." *Signal Processing*, *108*, 77–89.

Mielikainen, J. (2006). "LSB matching revisited." *IEEE Signal Processing Letters*, *13*(5), 285–287.

Milli, M., & Milli, M. (2023). Big Data and its Future. In *Data Science with Semantic Technologies* (pp. 27–43). CRC Press. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003310785-2/big-data-future-musa-milli-mehmet-milli

Saran, N., & Olcay, C. (2013). "İmge içine bilgi gizlemede kullanılan lsb yöntemlerinin karşılaştırması." *Cankaya University Journal of Science and Engineering*, *10*(1). https://dergipark.org.tr/en/pub/cankujse/issue/33158/368988

Sharp, T. (2001). "An Implementation of Key-Based Digital Signal Steganography." In I. S. Moskowitz (Ed.), *Information Hiding* (Vol. 2137, pp. 13–26). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45496-9_2

Swain, G. (2018). "High capacity image steganography using modified LSB substitution and PVD against pixel difference histogram analysis." *Security and Communication Networks*, *2018*. https://www.hindawi.com/journals/scn/2018/1505896/

Tian, J. (2003). "Reversible data embedding using a difference expansion." *IEEE Transactions on Circuits and Systems for Video Technology*, *13*(8), 890–896.

Volkhonskiy, D., Nazarov, I., & Burnaev, E. (2020). Steganographic generative adversarial networks. *Twelfth International Conference on Machine Vision (ICMV 2019)*, *11433*, 991–1005. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11433/114333M/Steganographic-generative-adversarial-networks/10.1117/12.2559429.short

Wang, C.-M., Wu, N.-I., Tsai, C.-S., & Hwang, M.-S. (2008). "A high quality steganographic method with pixel-value differencing and modulus function." *Journal of Systems and Software*, *81*(1), 150–158.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). "Image quality assessment: From error visibility to structural similarity." *IEEE Transactions on Image Processing*, *13*(4), 600–612.

Wu, D.-C., & Tsai, W.-H. (2003). "A steganographic method for images by pixel-value differencing." *Pattern Recognition Letters*, *24*(9–10), 1613–1626.

Yalman, Y. (2010). *Sayısal görüntüler için histogram temelli veri gizleme yöntemi ve uygulama yazılımı*. http://dspace.kocaeli.edu.tr:8080/xmlui/bitstream/handle/11493/15107/275801.pdf?sequence=1

Zi, H., Zhang, Q., Yang, J., & Kang, X. (2018). Steganography with convincing normal image from a joint generative adversarial framework. *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 526–532. https://ieeexplore.ieee.org/abstract/document/8659716/