Hiding text using the least significant bit technique to improve cover image in the steganography system

Estabraq Hussein Jasim Halboos, Abbas M. Albakry
Department of Computer Science, Iraqi Commission for Computers and Informatics (ICCI), Informatics Institute for Postgraduate Studies, Baghdad, Iraq

Article Info

Article history:

Received Jul 1, 2022
Revised Aug 5, 2022
Accepted Aug 23, 2022

Keywords:

Hiding text
Least significant bit
Steganography
Stenographic system

ABSTRACT

One of the highest priorities in the era of information technology is to achieve an accurate and effective system for hiding security data. One of the goals of steganography is imperceptability to intruder. So this paper work to increase the imperceptibility on image, which has weaknesses in previous studies, as well as to avoid statistical attacks such as chi-square. A method has been proposed that includes calculating the color contrasts in the homogeneous areas of the image and dividing them according to the color contrast and exploiting the data of pixels that have a high impact to embed on the two first and third bits of least significant bit (LSB) to increase the amount of embedded data, impact regions (IR) classify according to selected features extracted in advance by using the support vector machine (SVM) classifier. Work was done on standard images taken from a standard dataset (USC-SIPI) for two types of gray and color images. The results showed the worth of the proposed method through a high peak signal to noise ratio (PSNR) that reached 89.5 dB due to the distribution of data on pixels according to the proposed method.

Corresponding Author:

Estabraq Hussein Jasim Halboos
Iraqi Commission for Computers and Informatics (ICCI), Informatics Institute for Postgraduate Studies
Street AL-Nidal, 00964, Baghdad, Iraq
Email: estabraqhussein47@gmail.com

1. INTRODUCTION

The transfer of important and confidential data in our time via the Internet has become one of the biggest challenges due to the development in technology as well as the accumulation of experience over the years among those interested. There are three different basic methodologies to hide data and to ensure the security of data are encryption, steganography and watermark, thus they are no less important than the other [1]. Encryption is one of the basic ways to hide information because it works on a clear challenge in the form of cipher text, which may be simple at first glance, but in order to explain the complexity begins completely [2], [3]. As for concealment, it hides the text in certain media in an innocent way so as not to arouse suspicion during transmission. It is considered the best at the present time because of the great dependence on this type of security [4]. As for the third and last type, it is considered a common commercial use, which depends on the apparent or hidden watermark, which gives privacy to this method [5]. Steganography is one of the best secure communication techniques and it can be known as the science of invisible communication, which hides highly confidential information in an innocent-looking image cover [6], [7]. The main goal here is to design an algorithm to hide the information behind the given image and be more secure than the other. Steganography techniques work to hide the message from the ground up, so that the image carrying the message is unsuspecting by the third person [8]. Images are currently used because of their frequent circulation, as well as their frequent use in computers and internet. Digital data is the main carrier and networks are the rapid channels for transmission. Figure 1 illustrate the structure of steganography system and its concepts.

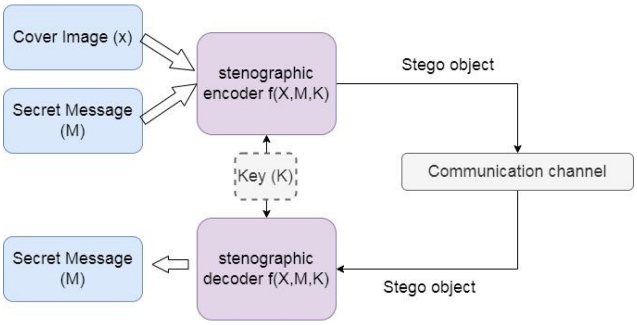


Figure 1. Steganography system structure and the main concepts within

As the image depicts, both cover file(X) and secret message(M) are fed into stenographic encoder as input. Stenographic encoder function, f(X, M, K) embeds the secret message into a cover file. Resulting Stego object looks very similar to your cover file, with no visible changes. This completes encoding. To retrieve the secret message, Stego object is fed into steganographic decoder. Hiding information includes two main steps [9]: hiding (embedding) and extraction. Hiding is the process of embedding the secret text inside the cover image in an innovative way so that the output is Stego Image, which goes to the other party, and the second step is to extract the confidential information from the image by the second party to extract the secret text. Steganography differs from cipher in that it preserves the content of an ambiguous message, while steganography keeps the existence of the message in an unobtrusive manner [10]. Also, there is often important information that works to maintain the confidentiality of information in a way that it is not revealed, as is the case in fingerprints and military and medical reports, and in some cases the information is unique, as in an eye and finger print.

As in Figure 1 the image resulting from the embedding process is called stego image, and then the system transmits the resulting image to the other party through communication channels and it is similar to the original image [11], and when it is received from the receiving party, it performs the process of extracting the text embedded inside the image with the help of the information available in the stego key and then it is produced image without secret data inside, as well as the secret data itself. The stego key is generated automatically and depending on the steganography method, without which the embedding cannot be decoded, and sent to the recipient to be used for extraction purposes [12]. One of the very common methods of hiding information is least significant bit (LSB) and this technique uses a less important part of the pixel to embed.

Secret messages can be hidden inside the cover image and in different types, including video files, text files, audio and images, communications networks, and even programs and in certain ways. As for hiding in images, it is considered the most thankful, because of its spread through the Internet at the present time [13], and because there is a wide scope for adding in this type of image. Hiding data in the cover image involves several methods that compete in increasing the imperceptibility of the main image, as well as the amount of payload capacity that the image can take. Steganography is an art that presents the cover as an innocent image that does not contain a secret, which is what distinguishes the work of the steganography technique. The method used in this research allows using the pixel value variant with its position and then choosing it to be a nominee for embedding and unfamiliar approach to the hacker or the intruder. The main contributions of the current work are as follows:

- It proposes a data hiding scheme employing reversible mathematical logic founded encoding. Secret bits are encrypted beforehand. The encoding parameters and the encryption key are secretly shared using public-key cryptography for high security.
- The scheme uses the order of the first and third bits LSB of the cover image pixel during encoding of the secret bits. Each secret bit takes place in the 1 of LSB and the next secret bit goes to 3 this process will have repeated for all image pixels with a certain condition.

The improvement and development of cover image will be via making the embedding of confidential data to the image in such a way that it cannot be seen or tracked, or rather, it is not even possible to feel it. The method of embedding two bits with the condition that the adjacent neighbor pixels are compatible is a guarantee to improve the cover of the image so that it is not possible to predict whether the image contains secret data or not, in addition to that if it is known that the image contains data, obtaining the data is almost impossible.

2. RELATED STUDY

Many different image-steganography methods utilizing LSB with different embedding techniques have been reviewed in the related work [14]-[19]. Cryptographic techniques are usually used in conjunction with steganography to input an extra-layer of security [20]-[22]. A clear analysis of image-steganography schemes along with commonly employed image quality metrics and measures can be seen in [23]-[26]. In the next paragraphs, we shall focus on the spatial domain LSB embedding_techniques that primarily employs some binary key and/or mathematical logic to encode the secret bits before hiding them in selected pixels.

Yedroudj *et al*. [27] proposed and designs for steganography system to hide secret data in a binary game, which highlights the power of hiding information in steganography and revealing in the announced network, while [28] relied on the design on LSB to hide secret message, but with a small percentage to hide the text bits in the image pixel and using encryption as well. Therefore, he designed a detection technique for some areas in the image to avoid the traditional methods, and on the other part, he reversed the embedding algorithm. Horng *et al*. [29] suggest a hiding method in the applied images by dividing the number by the actual value in the LSB, it is a method closer to encryption than to hiding, the method depends on three stages of division difference, substitution in the LSB, and embedding the digital change in the average value of the neighboring pixels, and considered a good method with more complexity. An unconventional method that was followed by [30] depends on the LSB attached to the image data and using genetic algorithms to find the optimal image for hiding, encouraging result obtained by this method. In order to hide data inside the image, various methods are used, and the simplest type of methods is the LSB. In this method, we exploit the less important bits of the image pixel to store data. The amount of data can be increased by increasing the number of bits exploited, but the imperceptibility will decrease in this case. This method is easy, but more vulnerable to attacks, especially statistical attacks, and chi-square analysis [31], [32].

A new technique was proposed by [33] using one channel for signal and two for embedding with a predetermined cycle to increase the robustness of the method used, the results of this method showed high payload capacity in addition to imperceptibility, and in return the method lacks a simpler key exchange. The LSB substitution technique presented by [34] which depends on the adaptive state of adding three bits per pixel in one way for the signal and two for embedding in the form of red, green, and blue (RGB) image, here the payload is relatively high for embedding bits in the central places and neglecting the areas that suffer from low contrast which increases its robustness. A successful method was suggested by [35] to embed the secret key in two channels of the green and blue colored image. The third channel (red) is in which the secret bits are included in the line spectral pair (LSP) method. In this method, the security level is high compared to other methods, while the amount of embedded data is small. The fact that the secret key security problem has been resolved was a challenge in itself. Mahato *et al*. [36] include studied a custom character to bit mapping to encode the secret text in the cover image LSBs. This is not standard encoding, and only the recipient learns about it. If exact encoding is not employed frequently, this makes it difficult for a third party to decode the secret text. Taha *et al*. [37] have used the modulus function and difference expansion (DE) to hide data in the spatial domain with increased capacity through leveraging smooth area in a cover image. Priya *et al*. [38] recently used DE method and interpolation based data hiding with non-sequential embedding into image pixels. In literature has to analyze the embedding to find the embedding ratio and start the comparison, to prove the efficiency of the system many images are used. When comparing the proposed system results with existing methods, the results were convincing as shown in Table 1.

Table 1. Benchmarking result of proposed method with existing method

Reference	Methods	PSNR/dB	Images
[39]	PVD	76	Baboon
[40]	DWT+ Hoffman	72	Baboon
[41]	Encryption +Vernam	66	Baboon
[42]	Bit shifting +LSB	81	Baboon
This work	IVR+ SVM	82	Baboon

The previous methods depend on one principle, which is the embedding in the places of the LSB, which is a rather easy to predict method, as in [43]. Many traditional methods take care of increasing security and this is at the expense of the amount of data embedded because a single pixel can accommodate a specified number of bits to maintain the imperceptibility of the image [44]. Choosing the right pixel to add is one of the most important things that can improve the cover of the image, which is considered a relatively difficult thing in the case of the imperceptibility of the image [45]. From this, a gap can be found to develop the proposed system to reach the best result.

3. PROPOSED METHOD

3.1. Steganography

Steganography is the art of hiding text in trusted media such image and consists of two main stages embedding and extracting. Data flow with these components controlled by steganalysisthat defines as the procedure used to detect any hiding of data in an image by differences in image size and pixel bit pattern, as shown in Figure 2. It is the art of detecting and presenting secret messages, and the aim of it is to know the suspect locations in the image or any other media, compare them, then identify them, if they are detected, then search if they are encrypted and return them as much as possible, and examine the result of the message being intercepted2 [18].

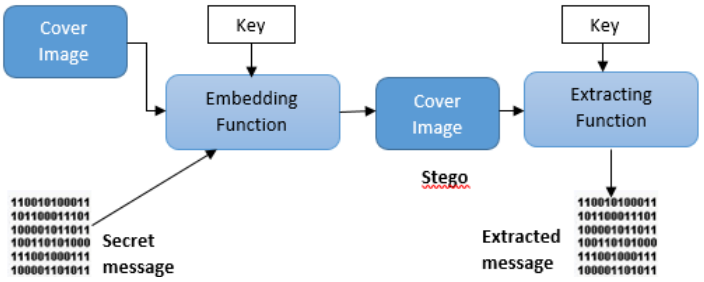


Figure 2. Steganographic system

The analysis starts from diagnosing the suspicious information in the image, and it is not often the image that contains hidden data. The available information is divided into smaller subtotals and processed statistically in advanced ways. Steganography is used in the following:

- Steganography is used as a method of hiding data in order to be easy to transmit and without supervision and to remove the fear of caution over messages being intercepted or tracked.
- Hidden information can be stored in websites, and this information is private, such as banking or military information, which is stored in the cover image, and which can only be accessed through the secret key, and therefore it is impossible to prove the existence of such information.
- Steganography is also used in watermarks, although watermarks do not often involve hiding information, but it is also known as the process of hiding information. steganography is intended to hide information and sometimes by means of watermarks that also use images and the usual methods of including data. Such a situation is found in the photos and videos.
- E-commerce is concerned with the use of data hiding, especially in ongoing electronic business transactions, and as they are in the password in the process uses to prove certain ownership, fingerprints also contain hidden information and also verifying trademarks increases the interests of data hiding.
- Data hiding can be used with the current methods of communication and its purpose is to exchange at a lower level and in return to attract government interest and national support. Security in most countries plays a key role in the country's politics and companies.

Hiding information is in several forms: encryption, watermark, and steganography. The steganography is one of the most important of these types, where the image loaded with secret data is innocent and does not raise any doubt for intruders, and the adoption of the image is that images are one of the most widely circulated media in the field of the internet. The resulting image after hiding the data is similar to the original image to the extent that its creator cannot be known, because hiding is in imperceptible way. Hiding data in the color gradients (or contrast) of the image allows for the possibility of manipulating the gradients in an amount that the human eye does not perceive, and therefore it is possible to embed in the contrast area of 30, which is not possible for the human eye to distinguish.

3.2. LSB technique

In the digital world, especially steganography, images are represented on the basis of color contrast in the manner of pixels, and each pixel consists of 8 bits, and the last four of these bits are called the LSB, which have an effect relative to the first bit by 1. Therefore, this feature or properties is used in images to hide data, and when using two bits, the inserted value is 3 and according to the formula 2n, which helps in adding more data and the technology that depends on LSB to replace the less important value in the pixels that cannot be noticed by human eye and since the LSB technology is widespread and well-known, so another technology is adopted to take more security and be unbreakable or easy to detect by intruders. The entry of data into the LSB (as proposed) is in the first and third bits in order to increase the security as well as to increase the complexity of the work in order to serve the increase of imperceptibility. However, the LSB approach has become mainly in inserting via the spatial domain and it is one of the very important techniques to add the most amount of information in the least number of pixels. With proposed method each secret bit take place in the 21 of LSB and next secret bit goes to 23 and this process will have repeated for all image pixels with certain condition will discuss in detail in the next section. The main stages of any LSB established steganography are: i) choose a type of cover if video, audio or image, ii) transform the secret message of video, audio, image or text into binary, iii) embed the secret binary bits into the LSB bits of the cover media intensity values to create the Stego media, iv) Stego media is transmitted to the recipient via the shared channel, and v) the receiver then extracts the LSB bits of stego cover and rebuilds the secret message.

4. METHOD

The proposed method uses one of the most popular media namely, images, as the cover. Hide text represents the secret text, which is transferred first to American standard code for information interchange (ASCII) code and then to binary in order to take into consideration the bits. And then hiding the text in the exact units of the image (pixels) to produce an image similar to the cover image [46], but contains hidden data that is not visible. Most of the previously used methods take the pixel LSB as the main target for the embedding and then make a specific change to suit the output. In our proposed method, we will study most of the attacks carried out by the hacker or the intruder and try to bypass the weaknesses and increase the gaps from which the stego image can be as innocent as possible [47].

Figure 3 shows the proposed method employs an eight-bit RGB cover image and uses a secret key while embedding. The stego key is enciphered and transmitted separately with the help of the public-key cryptography. The necessary stages in this method are as follows: i) encrypt the secret text/message with key encryption, ii) a cover image is chosen and split into image parts, iii) encrypted secret text/message is split into blocks for each part for encoding, and iv) encrypted bits are XORed with the first and third bits LSB of the cover image pixel. Proposed method classified into two stages first embedding secret bits into image by using impact varying technique will discuss in detail and second stage is extracting which inverse embedding procedure.

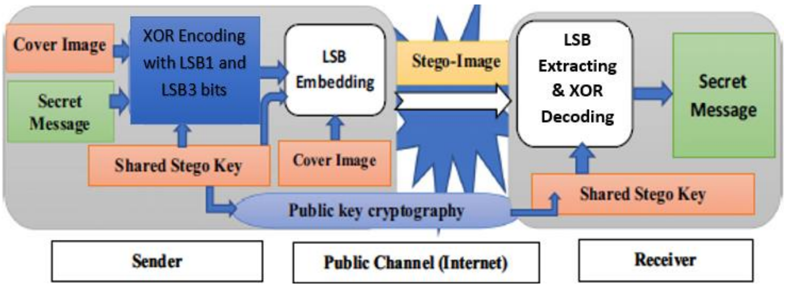


Figure 3. Block diagram of the proposed method

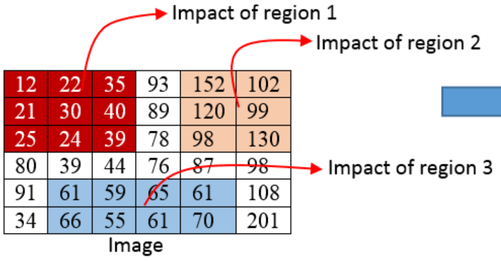
4.1. Embedding technique

Suggested image consist of 512×512 pixels each of 8-bits, to select which pixel candidate for embedding have to use random function. Using of random function also increase the security and consider the first step in security. Despite the random selection of the pixels, the output of the random function is not used directly, but rather is processed according to the proposed method in order to choose the pixel to carry the hidden data. Each image consists of groups of pixels that are close in color value and here comes the importance of the method, as shown in Figure 4.

The regions (sub-images) were selected according to the closeness of the values for each pixel in the image 512×512 or 1024×1024. There are many homogeneous regions that contain rich regions to add information. The main factor here is to find the difference between pixels' value of certain image and average impact of these value. In (1) shows how to calculate impact value for each region.

$$\text{impact region value} = \frac{1}{n}\sum x^n \tag{1}$$

Where x represent pixel value and n is the pixel number.



12	22	35	93	152	102
21	30	40	89	120	99
25	24	39	78	98	130
80	39	44	76	87	98
91	61	59	65	61	108
34	66	55	61	70	201

Figure 4. Construction of sub-image region and derived impact value

Vector of impact region considers as a features feed to the classifier to determine the best region to embed. Support vector machine (SVM) is used to classify the suggested region based on features extracted from the region itself, some features have high impact values such as variance and deviation that can be used to help classify issues for suitable embedding area. For security, random numbers are generated to choose pixels in the particular region before classification and then nominate pixels after classification, increasing the proposed method's security. Classifiers behave like cluster regions inside selected region in advance as shown in Figure 5.

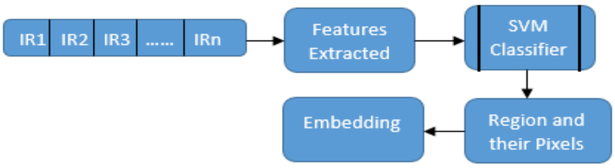


Figure 5. Classification of IR via SVM classifier

Regions are selected in advance to be appropriate for embedding and each is different from another to suitability for holding secret bits thus features like averaging, contrast varying neighbors, will extract and considered as labels for each IR then the classifier will arrange the priority for embedding, and choose if suitable or not. If the impact of certain region has low priority, then may not use according to classifier results. The classifier will train in 80% of the selected images in the dataset will result in a confusion matrix to illustrate the result and use 20% for testing mode. Dataset will have labeled according to standard results in the literature to train the system first. Procedure of collecting impact region (IR) value will scan all the image like filter and map the value into special vector to used later in embedding. Selecting pixel randomly in advance give the priority in embedding, due to candidate pixels less than pixels number of image then payload capacity will reduce to solve this problem two bits in LSB used for embedding as shown in Figure 6.

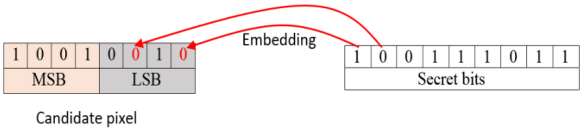


Figure 6. Embedding in 1st and 3rd LSB bits

Embedding here is to transfer secret bit into LSB pixel according to certain conditions and take place the bits in terms of replacement. Such as if the secret bit is 1 and the LSB bit is 0 then 1 value will locate instead of 0. The second secret bit will go directly to the 3rd LSB then the next pixel will come to take the same procedure and so on. During changing pixel value, the embedding will not recognize by human eyes due to human eyes can recognize more than 25 decimal value in pixel contrast while the maximum change in pixel value with the proposed method is 7 decimal value. The image resulting from the embedding process is an image that does not differ from the original one with a change in the less important bits in each pixel of the image, but for the naked eye it is indistinguishable. The difference between the image containing hidden data and the original image can be found only through some statistical treatments through which the proposed method is evaluated, and this will be discussing in detail later in the results section. Candidate pixels will manipulate according to embedding condition then will store again in stego image with different order and illustrated in (2) and (3).

$$R_i = \left(K_{i1}, K_{i2}, K_{i3}, \dots, K_{i(n-1)}\right) \tag{2}$$

$$K_{ij} = \left(P_{ij}^{(1)}, P_{ij}^{(2)}, P_{ij}^{(3)}, \dots, P_{ij}^{(n)}\right) \tag{3}$$

Where R consider the candidate pixel and i,j corresponding position, P is stego pixel contain hidden bit inside. In order for the process to be more contributing in terms of new ideas, the embedding is always in unknown regions, in a random sequence, and in a way that changes the pixel value as little as possible. All hidden data, of embedding, must be stored in stego key so that only the corresponding partner can see it, so that it can track the data embedded during the extraction process. The typical hierarchy of the method used is unpredictable and finding out where the data is hidden is almost impossible as shown in Figure 7.

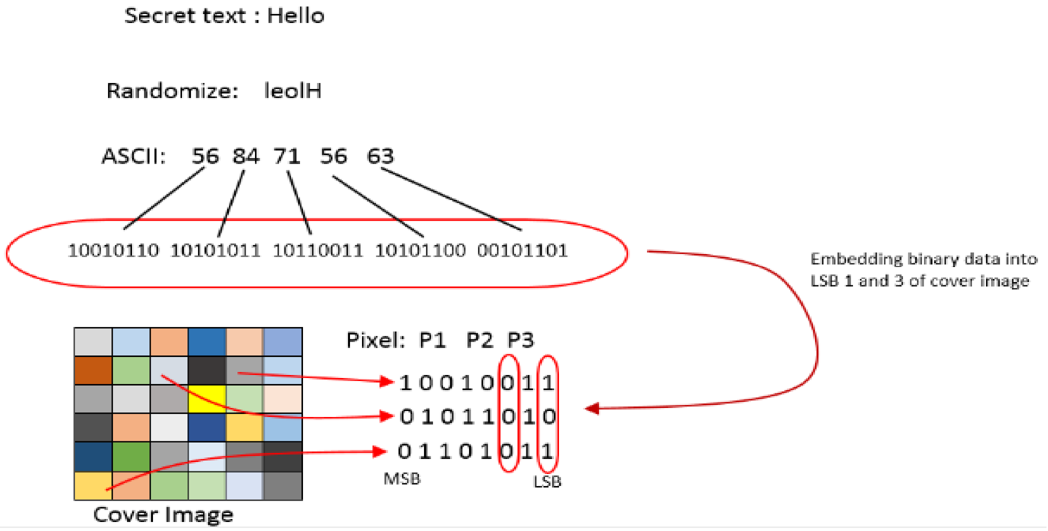


Figure 7. Process of embedding into cover image

In extracting process occur in the partner side when receive stego image then start to build algorithm of extraction according to the information in stego key. Extraction in general is the reverse process of embedding. Embedding consider cover image (original image) while extracting use stego image to process.

4.2. Classification

SVM is used as a famous and most efficient classifier in machine learning. It is used to select pixels and the candidate region for embedding, which are arranged in order of priority and importance. When regions are extracted and numbered in vectors, they are ready as features for entry into the classifier. Candidate regions for embedding, which are based on the amount of color contrast have been divided previously, there must be a way in order to prioritize the embedding and choose the appropriate pixel within a large set of candidate pixels. For this reason, we use SVM, which is of paramount importance, which receives data from vectors and classifies them according to the contrast and location of the pixels in relation to a specific area, thus producing vectors divided and arranged in the input sequence, so that we can then directly embed the secret data in the image as shown in Figure 8.

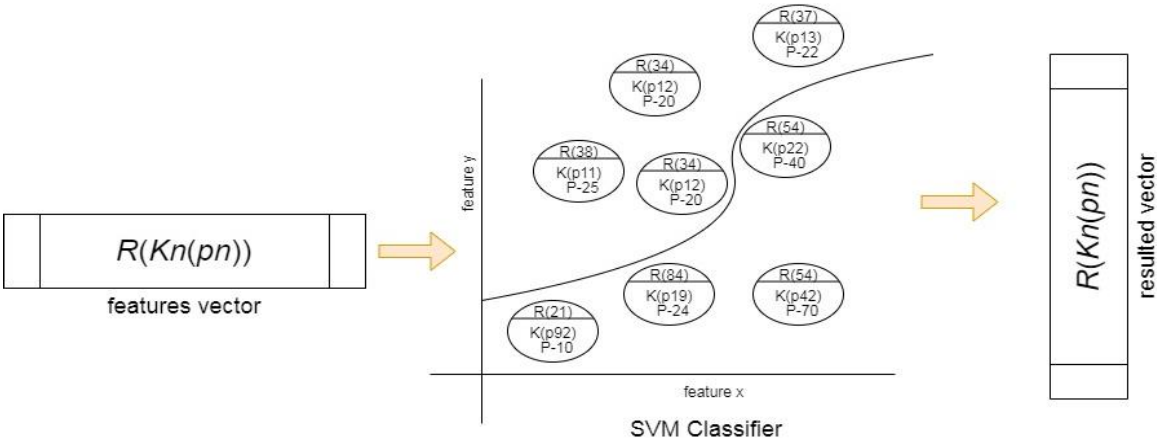


Figure 8. Classification of candidate pixels in features vector

4.3. Algorithm of embedding

{bring images from dataset then normalize it}
Step 1: prepare the image for embedding.
{process the spatial domain in term of image pixels to extract there features}
Step 2: calculate the regions according to pixels contrast

$$R^n = \sum_0^n Pi \times Pv$$

{manipulate subgroup of image to select bits of pixel to prepare for embedding}
Step 3: for all pixels in R Do

- Convert pixel value into digital value
- Find LSB1 and LSB3 for certain region
- Store in vector (VR) with label according to region properties

{prepare preprocessing issue before classification to find best features}
Step 4: extract features from VR
{recognize and select which image nominate for process}
Step 5: use SVM to classify the regions
{hide secret message in the image which is the main process of steganography}
Step 6: embed secret bits to Pixels in certain region in (LSB1 and LSB3) in same order.
{increase the capacity of the system by exploit all the image to hold more data}
Step 7: repeat for all R in image.

Embedding in LSB is very important due to is considered the main issue in embedding process of spatial domain. Many techniques suggested in literature use LSB each with different strategies, proposed method use priority of pixels' group in image and classify them to find the best position for embedding to get high imperceptibility and Figure 9 illustrate the LSB technique used in proposed method.

4.4. Extracting

Steganography is a process that takes place between the two parties of the sender and the recipient, where the sender embeds and all embedding operations are stored in the stego key and through this key the recipient on the second party can receive the stego image that contains the secret data and thus extract the data from it. There are two kinds of key explicit and implicit in our case explicit key were used that carry the process of embedding according to following algorithm.

Step 1: receive image g from sender.
Step 2: analyze the stego image g for extracting.
Step 3: use information from stego key to extract.
Step 4: extract regions R and P from vectors

$$R_i = \left(K_{i1}, K_{i2}, K_{i3}, \dots, K_{i(n-1)}\right)$$

$$K_{ij} = \left(P_{ij}^{(1)}, P_{ij}^{(2)}, P_{ij}^{(3)}, \dots, P_{ij}^{(n)}\right)$$

```
Step 5: for all $g(i,j)$ do
          Embed secret bit $S(i)$ into $g(i,j)_R$
Step 6: locate LSB 1&2 through $P(i,j)$
Step 7: extract secret bits
Step 8: repeat for vector $R$
```

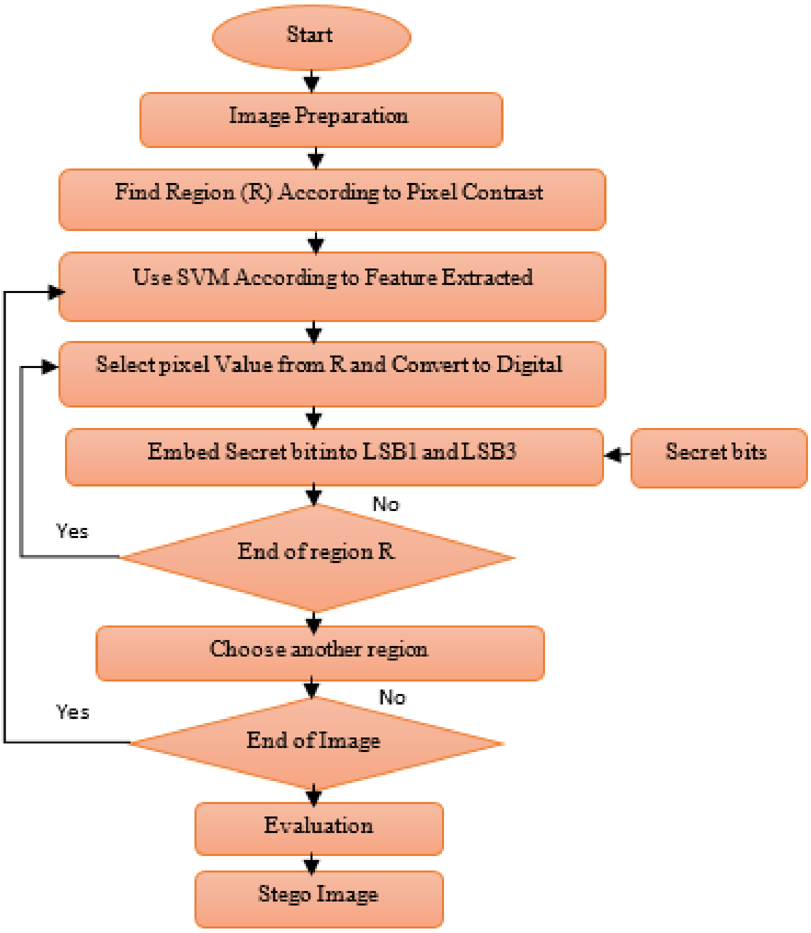

Start
Image Preparation
Find Region (R) According to Pixel Contrast
Use SVM According to Feature Extracted
Select pixel Value from R and Convert to Digital
Embed Secret bit into LSB1 and LSB3
Secret bits
No
Yes
End of region R
Choose another region
No
Yes
End of Image
Evaluation
Stego Image

Figure 9. Proposed technique based on LSB

5. RESULTS AND DISCUSSION

The quality of image processing is affected during processing, some information may be lost in the image, the usual methods of evaluation in this case are either objective or subjective. Objectivity depends on finding the difference between the resulting image and the original image and treating it statistically. The subjective method relies on human eye observation and judgment without reference to standards criteria. Many criteria used to evaluate such system will be discussed here according to its important.

5.1. Imperceptibility

It is a basic requirement in the process of hiding data in images, and it is necessary that the image not deteriorate after the embedding. In general, it is not possible to see the hidden data, and that the image appears innocent because it does not contain secret data. Peak signal to noise ratio (PSNR) [48] used to measure the imperceptibility of the system and can calculate as:

$$PSNR = 10.\,log_{10}\left(\frac{MAX_1^2}{MSE}\right) \tag{4}$$

and to identify mean square error (MSE) yields [49]:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \qquad (5)$$

where MAX represent the maximum pixel value in image with n, m consider dimensions of the image and I, K are the original and stego pixel (from cover and stego image). embedding 16,384 bytes to the cover image represents about 6.25% of the information that the image can hold in the LSB part. As the tables later show that the increase in capacity is inversely proportional to the value of the PSNR, and for a fair evaluation, the percentage of the embedding must be taken into consideration to standardize the embedding. Figure 10 shows the cover image and stego image after embedding 6.25% payload capacity.

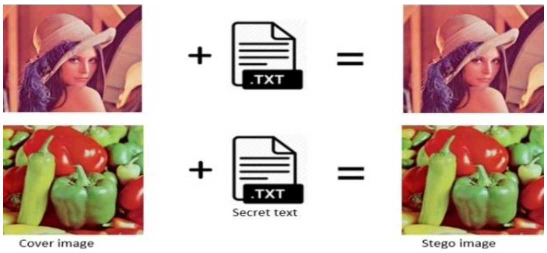


Figure 10. Cover image after embedding 6.25% payload capacity and stego image

With the proposed method, we used three amounts of payload capacity and three kinds of images derived from the standard dataset (the USC-SIPI image database) for better benchmarking. The three text sizes used in the proposed system are 16,384; 32,768; and 49,152 bytes represent 6.25%, 12.5%, and 18.75% respectively. This used image size 512×512 pixels, for a color image, gray image used the same procedure with a color image for embedding the difference in gray image one channel used for embedding and color image use the three channels RGB for embedding. Percentage of payload capacity calculated as (1 pixel = 8 bits) then 1/8 is 12.5% during embed one bit and (2 pixels = 16 bits) then 1/16 is 6.25% during embed one bit to the two pixels. In color images there is different case in steganography such as each color image has three type of mixed color RGB and embedding in this type of image will occur in three channels due to each contrast value comes from three different pixels one for red and one for green and one for blue mix of these pixels produce one color pixel. However, embedding in color image a bit higher imperceptibility than grey image as illustrated in Table 2. As we see in Table 2 when increasing the payload capacity then PSNR will decrease due to more data will be more detectable, so balancing the capacity with the quality of the image will be worthy. Also can be noticed that images with more varying contrast make more sub-images will appear then according to the proposed method more varying in pixel contrast can hold more secret bits. Peppers image have more regions and more varying color so it is more suitable to full these regions with data hiding.

Table 2. PSNR of standard images with different number of pixels (grey image)

Payload capacity	Percentage (%)	Lena image (dB)	Baboon image (dB)	Peppers image (dB)	Image /pixels
16,384	6.25	82.7	86.8	89.5	512×512
32,768	12.5	80.2	82.9	85.0	512×512
65,536	18.75	76.9	79,8	80.3	512×512
16,384	6.25	51.3	54.4	55.7	1024×1024
32,768	12.5	46.2	47.4	47.9	1024×1024
65,536	18.75	43.1	44.5	45.6	1024×1024

Figure 11 shows the relation between imperceptibility (quality of image) and amount of secret data can hide inside the image (payload capacity). If certain image carries more data, this will effect on the quality of the image and easy can recognized by human eye. In order to benchmark the results with the previous one, we must unify the embedding process, which is one of the main scales of the comparison process. In each study, a quantity of information is embedded to the image, which may be less or more than the researchers and other methods, so the results obtained do not express standard results, so the embedding quantity is standardized on the basis of the percentage of embedding embedding ratio (ER). ER calculated as changing in LSB of image pixel, and the percentage taken as 6.25% that’s mean imbedding one secret bit into two pixels, and 12.5% by embedding one secret bit into one pixel and so on. To illustrate the comparison of the proposed method with existing methods in terms of imperceptibility Figure 12 can depict the improvement of our method and how it worth.

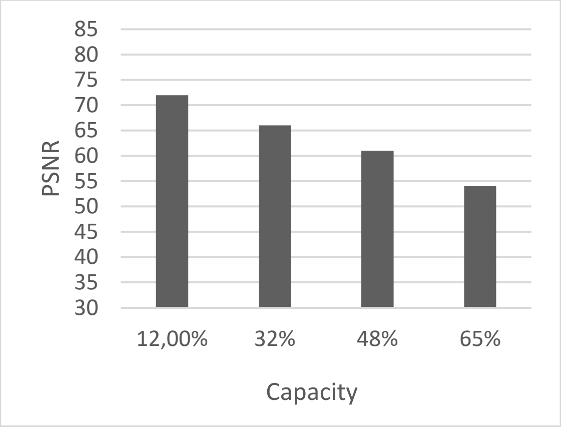


Figure 11. Relation between imperceptibility and capacity

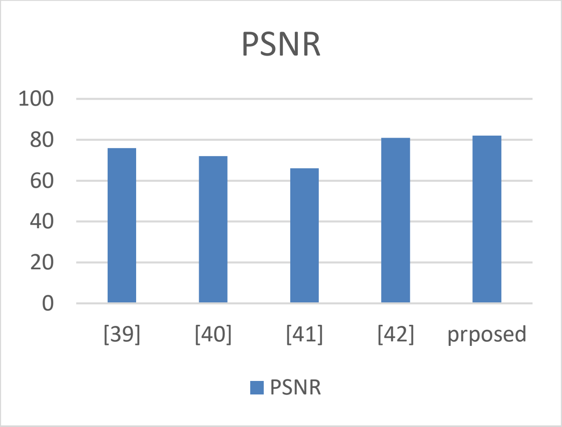


Figure 12. Comparison between the proposed method and existing methods

5.2. Chi-square attack (X2)

One of the objectives that have been taken into consideration in the proposed method is the second type of evaluation represented by chi-square, which is considered one of the most dangerous attacks used to detect the presence of data inside the image. A photo that contains data differs from a cover photo in that it is more vulnerable to attack. In this case, the attacker does not have the original image for comparison only the stego image, so he uses the arithmetic mean to calculate the frequencies and for each LSB bit and calculate their frequency in order to expect the presence of hidden data inside.

In this attack, which is called the statistical attack, it shows the possibility of including data in the image by repeating the LSB bits in the stego image, and as in the Figure 13 which displays the chi-square test, the x-axis represents the percentage of each image, and the y-axis shows the probability of including the secret message in the image. The first 3% has a probability of 0.065 if the units fulfil the function of each pixel. Most of the letters in the English language start with the same frequencies and the value of the repeated bits, so through this test, the frequency ratios are detected to know the embedding rate.

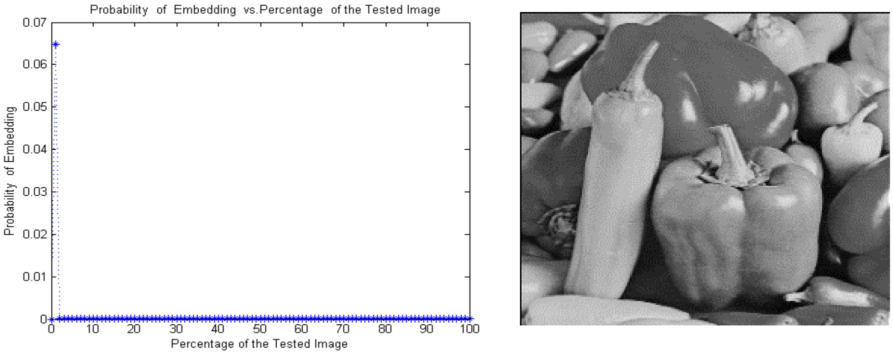


Figure 13. Detecting of chi-square for pepper image during proposed method

For embedding 18.75% to image 512×512 pixels' frequency of LSB stego pixels will be increased then embedding ratio will be clearly increases logarithmic due to the frequency of the rest pixels (not embedded) effect to statistical issue of chi-square equation.as shown in Figure 14. In short, chi-square reveals the security of images based on the embedding method. It also depends on the statistical distribution of the values of the pixels in the image and it is possible to obtain better results for the first part of the image because the beginning of the iterative calculation is inaccurate in this part of the image, so it is a suitable place to add where the frequencies are calculated according to the following equation.

$$X^2 = \sum \frac{(Observed - Expected)^2}{Expected} \qquad (6)$$

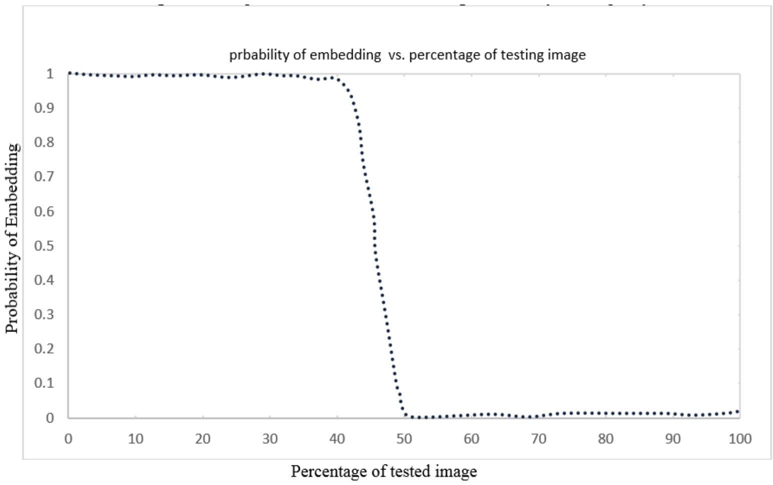


Figure 14. Chi-square behavior after embedding 18.75% payload capacity

6. CONCLUSION AND FUTURE WORK

Steganography is considered one of the most important arts of data hiding in more than one media. Data hiding depends on many variables that must be taken into consideration when designing the hiding algorithm, including the capacity, i.e. the amount of data to be stored in the image, the amount of security of the hidden data, in addition to imperceptibility, which by passing the picture without feeling that there is a secret inside it. The proposed method is important because of the attempt to avoid the attacks and to avoid the traditional methods based on LSB. The dataset used in this research is standard dataset called (USC-SIPI) including many images with different resolution for both gray and color images. Two of the most important evaluation criteria were used in this paper, which are imperceptibility and chi-square, and the results proved the worth of the method. In the future, it is possible to bypass the stage of the special domain and contribute to changing the variables that accompany the stage of the frequency domain and trying to integrate it with artificial intelligence (AI) techniques to choose the best cover image or choose the cover pixel itself.

REFERENCES

[1] R. Din, R. Bakar, S. Utama, J. Jasmis, and S. J. Elias, "The evaluation performance of letter-based technique on text steganography system," *Bulletin of Electrical Engineering and Informatics,* vol. 8, no. 1, pp. 291-297, Mar. 2019, doi: 10.11591/eei.v8i1.1440.

[2] M. H. Muhammad, H. S. Hussain, R. Din, H. Samad, and S. Utama, "Review on feature-based method performance in text steganography," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 1, pp. 427-433, Feb. 2021, doi: 10.11591/eei.v10i1.2508.

[3] O. Evsutin, A. Melman, A. E. -Latif, and A. Ahmed, "Overview of information hiding algorithms for ensuring security in IoT based cyber-physical systems," *in Security and Privacy Preserving for IoT and 5G Networks*," *Springer, Cham*, vol. 95, pp. 81–115, 2022, doi: 10.1007/978-3-030-85428-7_5.

[4] J. Liu, Y. Wang, Q. Han, and J. Gao, "A sensitive image encryption algorithm based on a higher-dimensional chaotic map and steganography," *International Journal of Bifurcation and Chaos,* vol. 32, no. 01, p. 2250004, 2022, doi: 10.1142/S0218127422500043.

[5] D. Pandey *et al*., "An integration of keyless encryption, steganography, and artificial intelligence for the secure transmission of stego images," in *Multidisciplinary Approach to Modern Digital Steganography*. IGI Global, 2021, pp. 211-234, doi: 10.4018/978-1-7998-7160-6.ch010.

[6] A. A. T. Rahem, M. Ismail, A. Idris, and A. D. Khaleel, " A comparative and analysis study of VANET routing protocols," *Journal of Theoretical & Applied Information Technology*,vol. 66, no. 3, pp. 691-698, 2014.

[7] N. J. Alhyani, O. K. Hamid, S. Y. Ali, and A. M. Ibrahim, "Efficient terrestrial digital video broadcasting receivers based OFDM techniques," *Przegląd Elektrotechniczny*, pp. 74-77, 2021, doi: 10.15199/48.2021.11.13.

[8] B. A. -E. -Atty, A. M. Iliyasu, H. Alaskar, A. E. -Latif, and A. Ahmed, "A robust quasi-quantum walks-based steganography protocol for secure transmission of images on cloud-based E-healthcare platforms," *Sensors*, vol. 20, no. 11, p. 3108, May 2020, doi: 10.3390/s20113108.

[9] A. A. A. E. -Latif, B. A. -E. Atty, and S. E. V. -Andraca, "A novel image steganography technique based on quantum substitution boxes," *Optics & Laser Technology*, vol. 116, pp. 92-102, Aug. 2019, doi: 10.1016/j.optlastec.2019.03.005.

[10] B. Karthikeyan, V. Abbinaiya, and T. Sumathi, "A novel approach in Steganography combining random key and substitution cipher," *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, 2019, pp. 673-678, doi: 10.1109/ICCS45141.2019.9065619.

[11] S. Pramanik, D. Samanta, S. K. Bandyopadhyay, and R. Ghosh, "A new combinational technique in image steganography," *International Journal of Information Security and Privacy (IJISP),* vol. 15, no. 3, pp. 48-64, 2021, doi: 10.4018/IJISP.2021070104.

[12] G. Maji, S. Mandal, and S. Sen, "Cover independent image steganography in spatial domain using higher order pixel bits," *Multimedia Tools and Applications*, vol. 80, no. 10, pp. 15977-16006, Feb. 2021, doi: 10.1007/s11042-020-10298-6.

[13] K. -S. Hsieh and C. -M. Wang, "Constructive image steganography using example-based weighted color transfer," *Journal of Information Security and Applications*, vol. 65, pp. 103-126, Mar 2022, doi: 10.1016/j.jisa.2022.103126.

[14] C. -K. Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," *Pattern* Recognit, vol. 37, no. 3, pp. 469–474, Mar. 2004, doi: 10.1016/j.patcog.2003.08.007.

[15] B. Datta, U. Mukherjee, and S. K. Bandyopadhyay, "LSB layer independent robust steganography using binary addition," *Procedia Comput* Science, vol. 85, pp. 425–432, 2016, doi: 10.1016/j.procs.2016.05.188.

[16] G. Maji and S. Mandal, "Secure and robust image steganography using a reference image as key," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 8, no. 7, pp. 2828–2839, May 2019.

[17] G. Maji, S. Mandal, S. Sen, and N. C. Debnath, "Dual image based LSB steganography," *2018 2nd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom)*, 2018, pp. 61-66, doi: 10.1109/SIGTELCOM.2018.8325806.

[18] P. Maniriho and T. Ahmad, " Information hiding scheme for digital images using difference expansion and modulus function," *Journal of King Saud University-Computer and Information* Sciences, vol. 31, no. 3, pp. 335–347, Jul. 2019.

[19] D. -C. Wua and W. -H. Tsai, "A steganographic method for images by pixel-value differencing," *Pattern Recognition* Letters, vol. 24, no. 9-10, pp. 1613–1626, Jun. 2003, doi: 10.1016/S0167-8655(02)00402-6.

[20] G. Maji, S. Mandal, N. C. Debnath, and S. Sen, "Pixel value difference based image steganography with one time pad encryption," *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, 2019, pp. 1358-1363.

[21] H. A. -Dmour and A. A. -Ani, "A steganography embedding method based on edge identification and xor coding," *Expert Systems with Applications*, vol. 46, pp. 293–306, Mar. 2016, doi: 10.1016/j.eswa.2015.10.024.

[22] Z. S. Younus and M. K. Hussain, "Image steganography using exploiting modification direction for compressed encrypted data," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 2951-2963, Jun. 2022, doi: 10.1016/j.jksuci.2019.04.008.

[23] N. M. Ganguly, G. Paul, S. K. Saha, and D. Burman, "A pvd based high capacity steganography algorithm with embedding in non-sequential position," *Multimedia Tools and Applications*, vol. 79, pp. 13449–13479, Jan. 2020, doi: 10.1007/s11042-019-08178-9.

[24] L. Singh, A. Singh, and P. Singh, "Secure data hiding techniques: a survey," *Multimedia Tools and Applications*, vol. 79, no. 23, pp. 15901–15921, 2020, doi: 10.1007/s11042-018-6407-5.

[25] C. A. Stanley, "Pairs of values and the chi-squared attack," *Department of Mathematics, Iowa State University*, 2005.

[26] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004, doi: 10.1109/TIP.2003.819861.

[27] M. Yedroudj, F. Comby, and M. Chaumont, "Steganography using a 3-player game," *Journal of Visual Communication and Image Representation*, vol. 72, p. 102910, Oct. 2020, doi: 10.1016/j.jvcir.2020.102910.

[28] O. Elharrouss, N. Almaadeed, and S. Al-Maadeed, "An image steganography approach based on k-least significant bits (k-LSB)," *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, 2020, pp. 131-135, doi: 10.1109/ICIoT48696.2020.9089566.

[29] J. -H. Horng, C. -C. Chang, and G. -L. Li, "Steganography using quotient value differencing and LSB substitution for AMBTC compressed images," in *IEEE Access*, vol. 8, pp. 129347-129358, 2020, doi: 10.1109/ACCESS.2020.3009232.

[30] J. C. T. Arroyo, J. A. Espadero, M. A. Ganas, R. F. Ardeña, R. N. Vilchez, and A. J. P. Delima, "An efficient least significant bit image steganography with secret writing and compression techniques," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 3, pp. 3280-3286, May-Jun. 2020, doi: 10.30534/ijatcse/2020/124932020.

[31] L. Mo, L. Zhu, J. Ma, D. Wang, H. Wang, "MDRSteg: large-capacity image steganography based on multi-scale dilated ResNet and combined chi-square distance loss," *Journal of Electronic Imaging*, vol. 30, no. 1, p. 013018, Feb. 2021.

[32] T. Manikandan, A. Muruganandham, R. Babuji, V. Nandalal, and J. M Iqbal, "Secure E-health using images steganography," *In Journal of Physics: Conference Series*, vol. 1917, no. 1, p. 012016, 2021.

[33] R. M. Rashid, B. Khalid Baker, O. F. Mohammad, and F. Y. H. Ahmed, "Information hiding in still image based on variable steganography technique to achieve high imperceptibility," *2021 IEEE 12th Control and System Graduate Research Colloquium (ICSGRC)*, 2021, pp. 171-176, doi: 10.1109/ICSGRC53186.2021.9515235.

[34] S. Rustad, A. Syukur, and P. N. Andono, "Inverted LSB image steganography using adaptive pattern to improve imperceptibility," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 6, pp. 3559-3568, Jun. 2022.

[35] A. I. H. Al-Jarah and J. L. O. Arjona, "Secret key steganography: improve security level of LSB algorithm," *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 2021, pp. 0215-0220, doi: 10.1109/UEMCON53757.2021.9666569.

[36] S. Mahato, D. K. Yadav, and D. A. Khan, "Personal characters to bits mapping using dot pattern character encoding scheme (DPCES)," *Journal of King Saud University - Computer and Information Sciences*, vol. 32, no. 2, pp. 197–207, Feb. 2020.

[37] N. A. Taha, A. A. Saffar, A. A. Abdullatif, and F. A. Abdullatif, "Image Steganography using dynamic threshold based on discrete cosine transform," *in Journal of Physics: Conference Series*, vol. 1879, no. 2, p. 022087, 2021.

[38] K. Priya, S. M. M. Roomi, P. U. Maheswari, and R. Suganya, "DWT based QR steganography," *in Journal of Physics: Conference* Series, vol. 1917, no. 1, p. 012020, 2021.

[39] D. R. I. M. Setiadi, "PSNR vs SSIM: imperceptibility quality assessment for image steganography," *Multimedia Tools and Applications*, vol. 80, no. 6, pp. 8423-8444, Nov. 2021, doi: 10.1007/s11042-020-10035-z.

[40] C. A. Sari, G. Ardiansyah, De R. I. M. Setiadi, and E. H. Rachmawanto, "An improved security and message capacity using AES and Huffman coding on image steganography," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 17, no. 5, pp. 2400-2409, Oct. 2019, doi: 10.12928/TELKOMNIKA.v17i5.9570.

[41] E. Z. Astuti, D. R. I. M. Setiadi, E. H. Rachmawanto, C. A. Sari, and M. K. Sarker, "LSB-based bit flipping methods for color image steganography," *in Journal of Physics: Conference Series*, vol. 1501, no. 1, p. 012019, 2020.

[42] S. Solak and U. Altinişik, "A new approach for steganography: bit shifting operation of encrypted data in LSB (SED-LSB) ," *Bilişim Teknolojileri Dergisi*, vol. 12, no. 1, pp. 75-81, 2020, doi: 10.17671/gazibtd.435437.

[43] S. Kaur, S. Singh, M. Kaur, and H. N. Lee, "A systematic review of computational image steganography approaches," *Archives of Computational Methods in Engineering*, pp. 1-23, Jun. 2022, doi: 10.1007/s11831-022-09749-0.

[44] Y. Bhavani, P. Kamakshi, E. K. Sri, Y. S. Sai, "A survey on image steganography techniques using least significant bit," *In Intelligent Data Communication Technologies and Internet of Things,* Springer, Singapore, vol. 101, pp. 281-290, Feb. 2022, doi: 10.1007/978-981-16-7610-9_20.

[45] B. L. Sirisha B. C. Mohan, "Review on spatial domain image steganography techniques," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 4, no. 6, pp. 1873-1883, Sep. 2021, doi: 10.1080/09720529.2021.1962025.

[46] P. C. Mandal, I. Mukherjee, B. N. Chatterji, "High capacity reversible and secured data hiding in images using interpolation and difference expansion technique," *Multimedia Tools and Applications*, vol. 80, pp. 3623–3644, 2021, doi: 10.1007/s11042-020-09341-3.

[47] Y. Yang, L. Zha, Z. Zhang, and J. Wen, "An overview of text steganalysis in the international conference on image," *Vision and Intelligent Systems (ICIVIS 2021)*, Springer, Singapore, vol. 813, pp. 933-943, Mar. 2022, doi: 10.1007/978-981-16-6963-7_82.

[48] J. Qin, Y. Luo, X. Xiang, Y. Tan, and H. Huang, "Coverless Image steganography: a survey," in *IEEE Access*, vol. 7, pp. 171372-171394, 2019, doi: 10.1109/ACCESS.2019.2955452.

[49] E. J. Kusuma, C. A. Sari, E. H. Rachmawanto, D. R. I. M. Setiadi, "A combination of inverted LSB, RSA, and arnold transformation to get secure and imperceptible image steganography," *Journal Of ICT Research & Applications*, vol. 12, no. 2, pp. 103-122, 2018, doi: 10.5614/itbj.ict.res.appl.2018.12.2.1.

BIOGRAPHIES OF AUTHORS

Estabraq Hussein Jasim Halboos currently is a Master's researcher at Informatics Institute for Postgraduate Studies (IIPS), Iraqi Commission for Computers and Informatics (ICCI), Baghdad, Iraq. He received her Diploma in computer science from Informatics Institute for Postgraduate Studies at the Iraqi Commission for Computers and Informatics, Iraq, in 2019. He received a B.Sc. in computer science from the University of Technology, Baghdad, Iraq, in 2010. He researches interests include artificial intelligence, machine learning, bioinformatics, and cyber security. She can be contacted at email: ms202030596@iips.icci.edu.iq and estabraqhussein47@gmail.com.

Prof. Abbas M. Al-Bakry he is currently a senior lecturer at University of Information Technology and Communications. Research interest: software agents, internet, web sites, artificial intelligence, bar codes, biomedical MRI, brain, data handling, data mining, face recognition, feature extraction, image classification, image denoising, image segmentation, information filtering, intelligent transportation systems, learning (artificial intelligence), matrix algebra, medical image processing, multi-agent systems, object recognition, query processing, radiofrequency identification, real-time systems, and road vehicles. He can be contacted at email: abbasm.albakry@uoitc.edu.iq.