



Image Steganography: Secret Data Detection in Digital Images using Least Significant Bit Method

Ong Chong Meng¹, Shahrum Shah Abdullah^{1*}

¹ Department of Electronic System Engineering, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:

Received 18 March 2020

Received in revised form 27 August 2020

Accepted 18 September 2020

Available online 29 September 2020

Keywords:

Image steganography; least significant bit; advanced encryption standard; secret messages

ABSTRACT

Nowadays, the number of data breaches increases surprisingly with the advancement of digital information storage. To increase the safety level of confidential data stored in the digital world, image steganography was proposed to prevent information leakage to intruders. Image steganography is the process of concealing secret information within a digital image. Most image steganography methods have compatibility, security and imperceptibility challenges: limited secret data embedded, brute force secret data extraction and suspicious and perceptible stego image generated. This research had proposed a novel image steganography method to hide confidential information inside a digital image using the Least Significant Bit (LSB) substitution. The usage of a compressed archive file for accepting multiple secret message was to be encrypted. A strong Advanced Encryption Standard (AES) symmetric encryption was combined with the Secure Hash Algorithm-256 (SHA-256) hashing algorithm to provide secret key protection to increase security. Random substitution in LSB increased the imperceptibility of the stego image generated. Secret messages were hidden in the stego image where human eyes cannot detect the difference compared to the cover image. The imperceptibility of the proposed method was evaluated in terms of parameters Peak Sound to Noise Ratio (PSNR) and Structural Similarity Index (SSIM). The overall PSNR results obtained were above 50 dB, where the minimum acceptable value was only 30 dB. Visual comparison and histogram analysis indicated that the changes in the visual structure of the image were very tiny, which will not be more than 0.02 %, as proven in SSIM results. In conclusion, the proposed LSB image steganography is a secure steganography method that accepts multiple types of secret messages to generate an imperceptible stego image with the secret key protection using AES.

1. Introduction

The presence of the Internet had connected all people around the world. The development of the Internet had led the world into the digital century, where most of the information and data were stored in digital form. Terabytes of data can be stored safely in a small hard drive, compared to the

* Corresponding author.

E-mail address: shahrum@utm.my

paper documents containing the same amount of information that will require a storage room. The cloud storage services to store digital data had increased the accessibility where people can access information anywhere using the Internet. However, digital storage using the Internet had created an opportunity for hackers to steal personal or confidential information besides bringing ease to people.

Therefore, the practice to prevent confidential data from being exposed to anyone is essential. The outgoing information on the Internet to remain secret can reduce the chance of data exposed to intruders even data breaches occurred. One of the methods to hide and conceal confidential information stored in the digital world is image steganography [1]. Steganography is the process of concealing secret information within other information media such as audio, video, text and image [2]. Image steganography is the type of steganography using the digital image as the medium to hide and conceal important information. Steganography is the art of concealing secret information within other information media such as audio, video, text and image [2]. The term "Steganography" comes from the combination of two Greek words, "Steganos" and "Graphein", bring the meaning of concealed writing. The word "Steganos" brings the meaning covered or concealed, while the word "Graphein" means writing or drawing [3]. Steganography is the standard method used to conceal the existence of secret information in a cover medium. Payload capacity is the first valuable measure to determine available space and amount to insert the secret message in the image steganography [4]. The product of the secret data embedded into the cover media is called the stego-object. The stego-object can be used for data transmission on the Internet securely with high confidentiality [5].

Imperceptibility is the third valuable measurement to define the effectiveness of steganography. Peak Sound to Noise Ratio (PSNR) and Structural Similarity Index (SSIM) are the standard index in the imperceptibility test used to identify the ability of stego image to be imperceptible to human eyes [6]. Thus, achieving a high value of imperceptibility is the essential consideration in determining the method of steganography.

The goal of the image steganography is to conceal the secret message (payload) into a digital image such that the presence of the payload will be imperceptible to intruders. The Least Significant Bit (LSB) substitution method is the simplest terminology in image steganography to archive this objective. Each LSB in the pixels of the image will be replaced by the secret message in binary format and produced an image called a stego image. The stego image as the product of image steganography is identical to the original image as the changes occurred only at the LSB. Thus, the maximum embedding capacity for the LSB method is only 3 bits for each pixel in the colour image [7].

This research focuses on concealing and detecting confidential information in a digital image. The LSB substitution is the image steganography method used to hide a secret message in the digital image. The cover image used in image steganography must be resistant to data degradation during transmissions such as .bmp or .png format. The .png format was picked as the type of image format used because of its common use in data transmission. Secret key encryption with the aid of Advanced Encryption Standard (AES) was asserted in the embedding process to increase the privacy and security of image steganography. The same secret key was used to decrypt and retrieve the secret message. The format of the secret message embedded will be studied along with the compression of digital data. The proposed image steganography aimed to apply all types of secret information embedded in the cover image. The imperceptibility of the stego image produced will then be tested with the value of PSNR and SSIM.

2. Literature Review

2.1 Concept of Steganography

Steganography is the standard method used to conceal the existence of secret information in a cover medium [15]. The product of the secret data embedded into the cover media is called the stego-object. The stego-object can be used for data transmission on the internet securely with high confidentiality [8]. The ability of steganography to conceal personal data had reduced the probability of confidential data leakage to intruders. Secret key cryptography and the hash function will be combined in the encryption and decryption of steganography to increase the safety level in this research. Figure 1 shows the encryption and decryption process of the steganography [9].

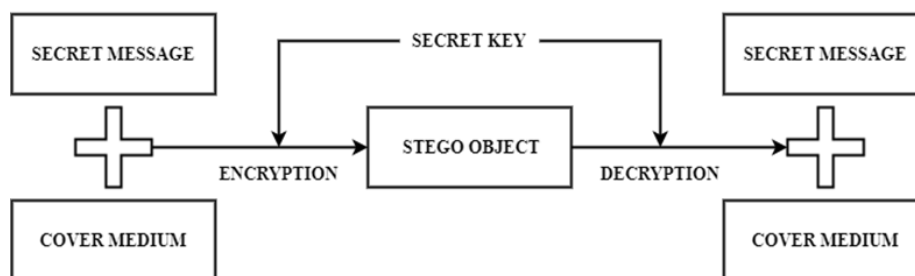


Fig. 1. Process of steganography

In steganography, the user will encrypt the secret message with the cover medium and insert the secret key to form a product named stego object. Stego objects will have no difference compared to the cover medium with the human visual system. The stego object will have different names depending on the cover medium. If the cover medium is a digital image, it will be a stego-image; the same goes for the video or sound as the cover media. The authorized user will receive the stego-object from the sender. The same secret key during encryption will be required to decrypt and extract the secret message from the stego-object. The retrieval algorithm of steganography will be the same as the embedding algorithm.

2.2 Types of Steganography

Steganography is divided into text steganography, image steganography, audio steganography and video steganography. Figure 2 shows the types of steganography in general [10].

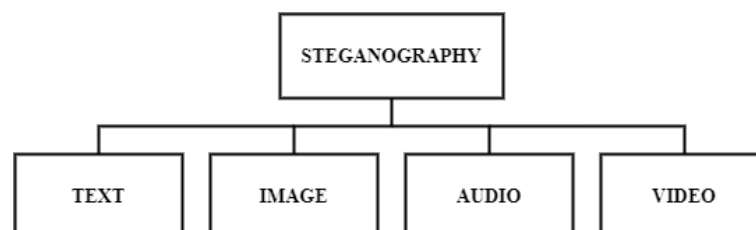


Fig. 2. Types of steganography

The classification of steganography types is based on the data type of the cover medium used. For example, some text steganography will use ciphertext as cover media to hide secret information because of its unreadable properties. However, the most common text steganography apply the technique to hide the secret message in every N^{th} character inside the cover text.

Audio will be the cover medium used to conceal the secret message in audio steganography. The masking technique hides a secret message in the cover audio to avoid detection from the human hearing system. Video steganography uses a video as the cover medium to hide confidential data from being detected. Its advantage of enormous capacity can hide the large size of secret messages.

2.3 Image Steganography

In image steganography, the confidential information will be embedded into the cover image based on the embedding algorithms [14]. A lossless digital image format should be considered to avoid data degradation during transmission. The secret key is essential in the encryption and decryption process for safety purposes. The stego image encrypted can be sent in data transmission because of its ability to deceive hackers and intruders. It will be an ordinary image identical to the cover image without doing any further analysis. The secret message in the stego image will remain confidential even if it is exposed to the hackers when a secret key is required in decryption to retrieve the secret information. Figure 3 shows the process of encryption and decryption in image steganography [11].

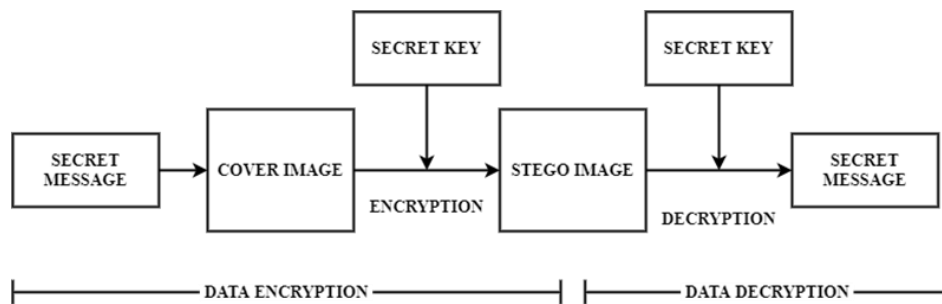


Fig. 3. Process of image steganography

2.4 Least Significant Bits (LSB) Substitution

The LSB substitution is the technique of image steganography used in this research. This technique will replace every LSB in the digital image with the secret data targeted to be hidden [9]. The binary data of each colour channel in digital images' pixel is where the LSB substitution method takes action. The secret data that wished to be hidden will be converted into binary values, and it will replace every single LSB of the colour channels in the pixels. The stego image produced will remain the same picture as the cover image since the data changes are too small to be detected by the human visual system [12]. Figure 4 shows the mechanism of the LSB substitution method in image steganography [13].

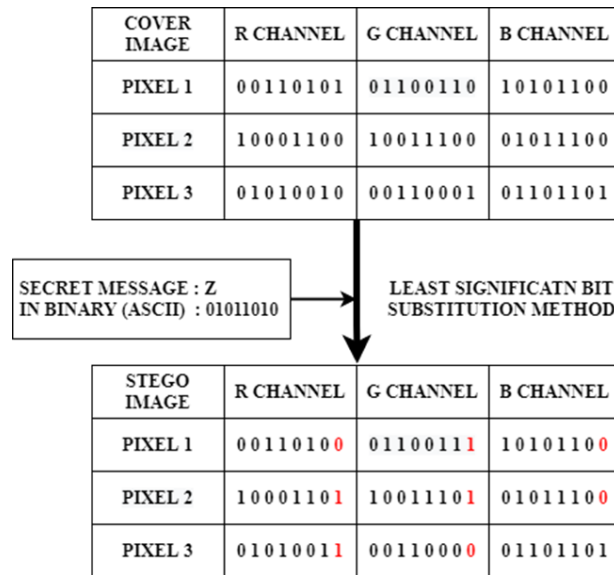


Fig. 4. Concept of LSB substitution method

3. Methodology

3.1 Data Preparation

To investigate the effectiveness of the proposed image steganography method, the cover image and secret message used were prepared compared to the section data evaluation. The size of cover images used was similar to other researchers’ standards. The secret message used was prepared in digital image format according to the percentage of payload capacity embedded. In this research, the cover image used in the proposed image steganography method for comparison was the standard image for image processing: Lena and Lake in png format. The cover image used was a 24-bit colour image having the standard size of 512 x 512 pixels. The .png format was chosen to be the format of the cover image because of its lossless properties, capable of avoiding data degradation during the transmission process. Figure 5 shows the cover image used in the proposed image steganography method.

Secret messages can be encrypted in any type of format in the proposed image steganography method due to the application of compression to an archive file. The entire secret message was compressed into a zipped file. The zipped file is the only type of information needed to be encrypted by symmetric key cryptography before embedding into the digital image. Thus, multiple files were applicable in the proposed image steganography to be embedded in the digital image. However, all the secret messages prepared to be colour images to be encrypted in the digital image for data comparison.

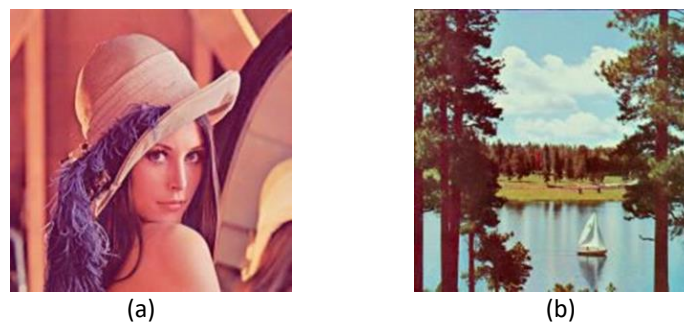


Fig. 5. Cover images used (a) Lena (b) Lake

3.2 Data Encryption

In the process of encryption in the proposed image steganography method, there were three inputs: the secret key, the secret message and the cover image. Figure 6 below shows the encryption algorithm of the proposed image steganography method. The secret key inserted in the encryption algorithm was passed to the Secure Hash Algorithm-256 (SHA-256) to generate a 256-bit hash value. The hash value generated was used for three purposes: the initialisation vector for Advanced Encryption Standard-128 (AES-128), the secret key for AES-128 and the generation of the random number that acted as an encryption indicator in this algorithm. The first 128 bits of the hash value was fed to the AES-128 as an initialisation vector to encrypt the secret message. A secret key was needed to encrypt the message in AES-128. Thus, the second 128 bits of the hash value were the secret key for secret message encryption. Besides that, the 256-bit hash value was then used as the seed to generate a random number by including the total numbers of encodable bits of the digital image. The random number in binary format represented the numbers' position for encryption, which each bit in the data representing one pixel in the digital image.

AES-128 in the encryption algorithm encrypts the secret message into ciphertext format no matter it is in plain text or file format. Multiples file encryption was applied in the proposed image steganography method because all the files inserted will be compressed into a zip archive. The zip archive will then will be encrypted in AES-128 with the initialization vector and secret key from the hash value generated by the SHA-256. The secret message in ciphertext format was now ready to be encrypted in the cover image using the LSB substitution method.

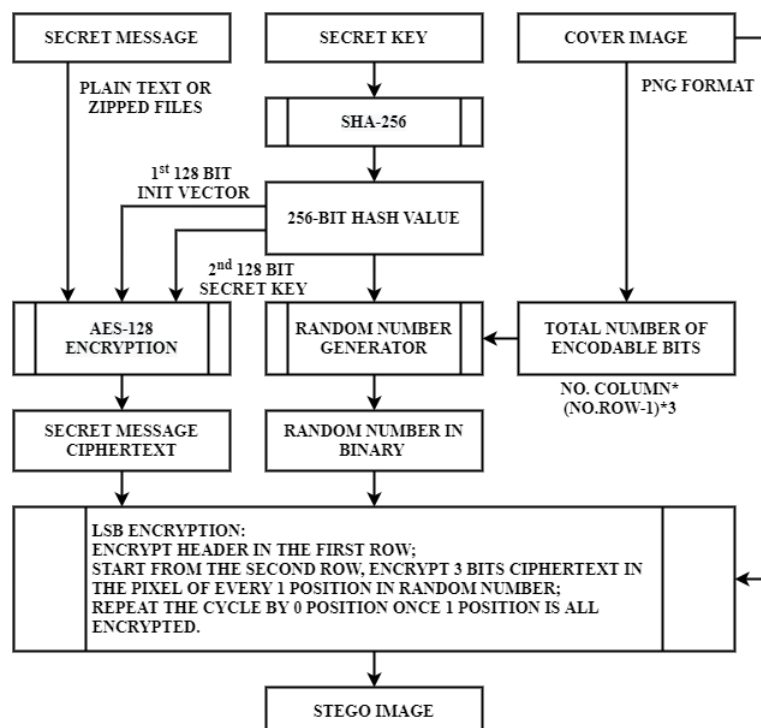


Fig. 6. Encryption algorithm of proposed image steganography

The encryption in image steganography started in the first row of the cover image. The header containing the length of the ciphertext message and zipped files header were encrypted in the first row of the cover image to instruct the start and end of the LSB substitution. Then LSB substitution continues at the second row of the image to replace 3 bits of ciphertext secret message in the pixels in the one (1) position, referring to the random number generated. The data replacement continues

until every pixel in the one position was substituted. Then the cycle will repeat with the pixels in the zero (0) position indicated by the random number. A stego image encrypted with the secret message in ciphertext format was then produced.

The proposed image steganography had improved most of the weakness of the LSB substitution method. A secret key was needed to encrypt the secret message so as the decryption process. Unauthorised intruders without the secret key cannot decrypt and retrieve back the secret message hiding in the stego image. Besides that, the random data encryption using the random number generated by the secret key had significantly decreased the possibility of image steganography being detected. Every secret key with different cover images will generate a different encryption algorithm in the proposed image steganography. Therefore, hackers that do not understand the mechanism of proposed image steganography were impossible to extract the hidden secret message out of the stego image. Eventhough the hackers or intruders can obtain the hidden secret message in the stego image; encrypting secret message into ciphertext format using the AES-128 makes it impossible for hackers to read the secret message in ciphertext format. Moreover, the proposed steganography using the LSB substitution method had minimized the changes of data in the stego image, making the stego image safe and less suspicious from being exposed.

3.3 Data Evaluation

To investigate the performance of the proposed image steganography method, the stego images produced were evaluated by comparing them to the original cover image. Evaluation in the human visual system and digital analysis should confirm that the proposed image steganography was secure and imperceptible. Image comparison and imperceptibility testing between the cover image and stego image generated was the evaluation, which can display the ability of the proposed image steganography method in achieving objectives.

3.4 Image Comparison

Both the cover image and stego image were compared by the human visual system and digital histogram analysis. The differences between the cover image and the stego image should be too small to be detected by visual comparison. Furthermore, histogram analysis was made between the cover image and stego image in terms of the pixel intensities separated into three primary RGB colour channels (red, green and blue). Comparison was made between the histogram of the cover image and the stego image to differentiate the colour changes.

3.5 Data Decryption

In the process of decryption in the proposed image steganography method, there were two inputs: the secret key and the stego image. Figure 7 below shows the decryption algorithm of the proposed image steganography method. The specific secret key during the encryption process was inserted once again in the decryption process. The secret key was passed to the SHA-256 to generate the same 256-bit hash value. The 128-bit initialization vector and the 128-bit secret key for AES-128 obtained from the hash value was the same compared to the encryption process. Only the same initialization vector and secret key can decrypt the secret message in ciphertext format extracted from the stego image. The random number generated by the 256-bit hash value and the total number of decodable bits was the same as the size of the cover image, and the stego image is the same.

However, the random number in the binary format now indicates the numbers' position for decryption, which each bit in the data representing one pixel in the digital image.

The decryption in image steganography started in the first row of the cover image. The header in the first row containing the length of the ciphertext message was read to control the extraction of the LSB of data. The decryption continues at the second row of the image to retrieve 3 bits of ciphertext secret message in the pixels in the one (1) position, referring to the random number generated. The data extraction continues until every pixel in the one (1) position is retrieved. Then the cycle will repeat with the pixels in the zero (0) position indicated by the random number. A secret message in ciphertext format is then decrypted from the stego image.

The secret message in ciphertext format will be fed to AES-128 to decrypt back to a secret readable message. The same initialization vector and secret key from the 256-bit hash value enable AES-128 to retrieve the original secret message. The zip file header containing multiple compressed files was read to extract the secret message back. Finally, the secret message encrypted in the stego image will be retrieved.

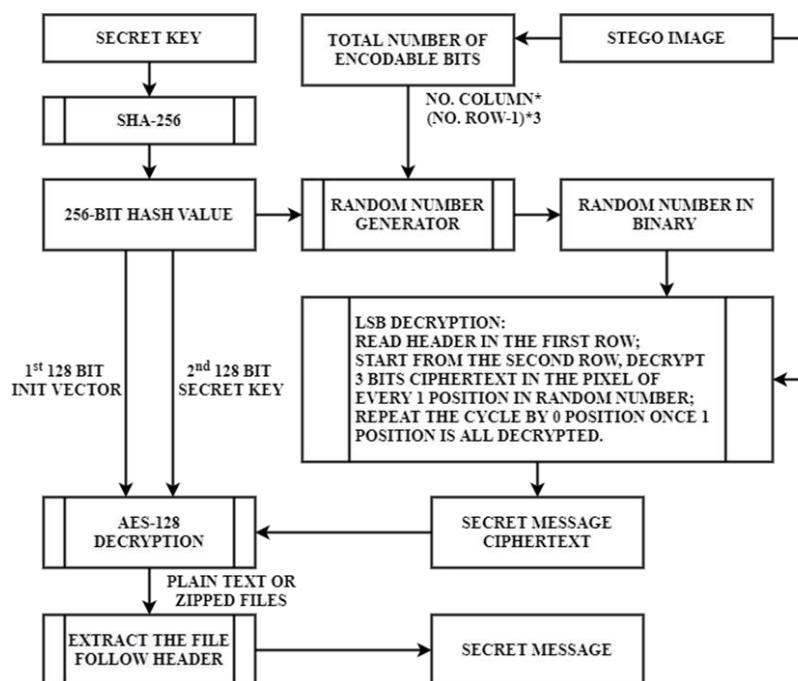


Fig. 7. Decryption algorithm of proposed image steganography

4. Results and Discussion

4.1 Visual Comparison

There are two cover images: Lena and Lake involved in the visual comparison with the stego image produced with 100 % of payload capacity in the proposed image steganography. Figure 8 shows the comparison between the Lena cover image with its stego image, while Figure 9 shows the comparison between the Lake cover image with its stego image.

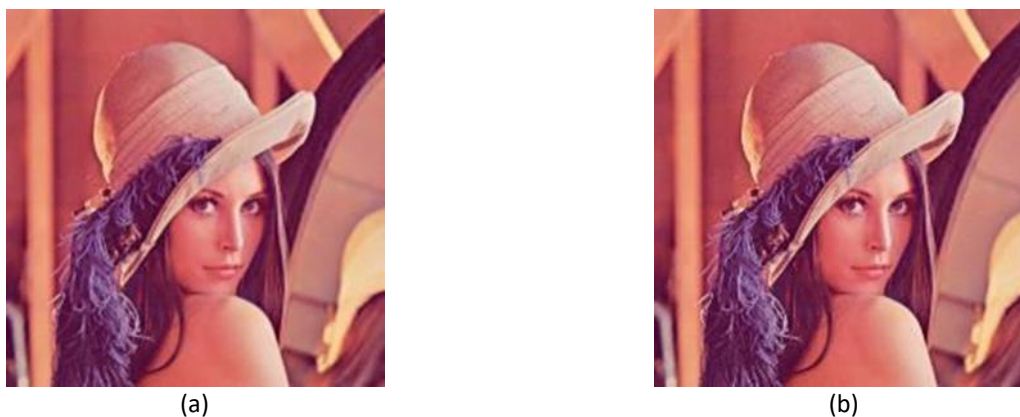


Fig. 8. Visual comparison between Lena (a) Cover image (b) Stego image

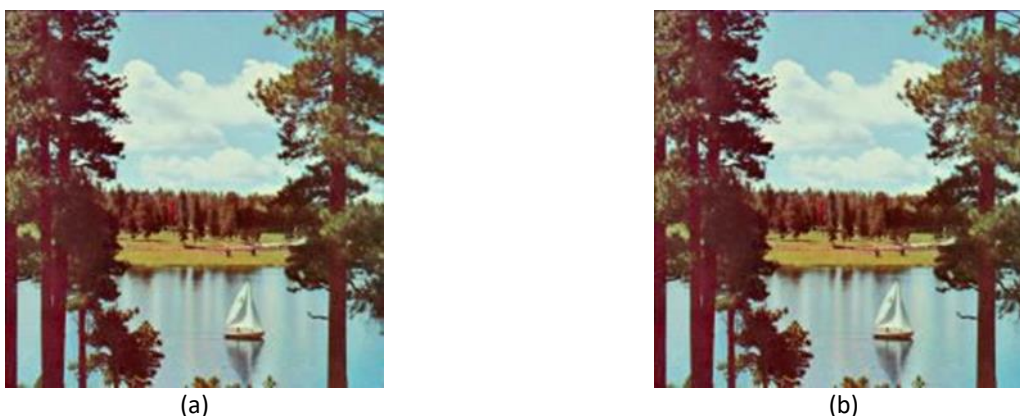


Fig. 9. Visual comparison between Lake (a) Cover image (b) Stego image

No difference can be found in the visual comparison between the cover image in the stego image in the Lena and Lake sets. Both images look exactly the same, which the proposed image steganography was imperceptible to the human visual system. The changes in the image were minimal until it was invisible to the human visual system.

4.2 Histogram Analysis

Histogram analysis in three colour channels was evaluated between the cover image and the stego image in the Lena and Lake sets. Comparison was made between the colour channels of the cover image and the stego image. The red colour indicates the histogram of the cover image, while the blue colour indicates the histogram of the stego image. Figure 10 shows the three colour channels histogram, including the comparison of the cover image and stego image in the Lena Set.

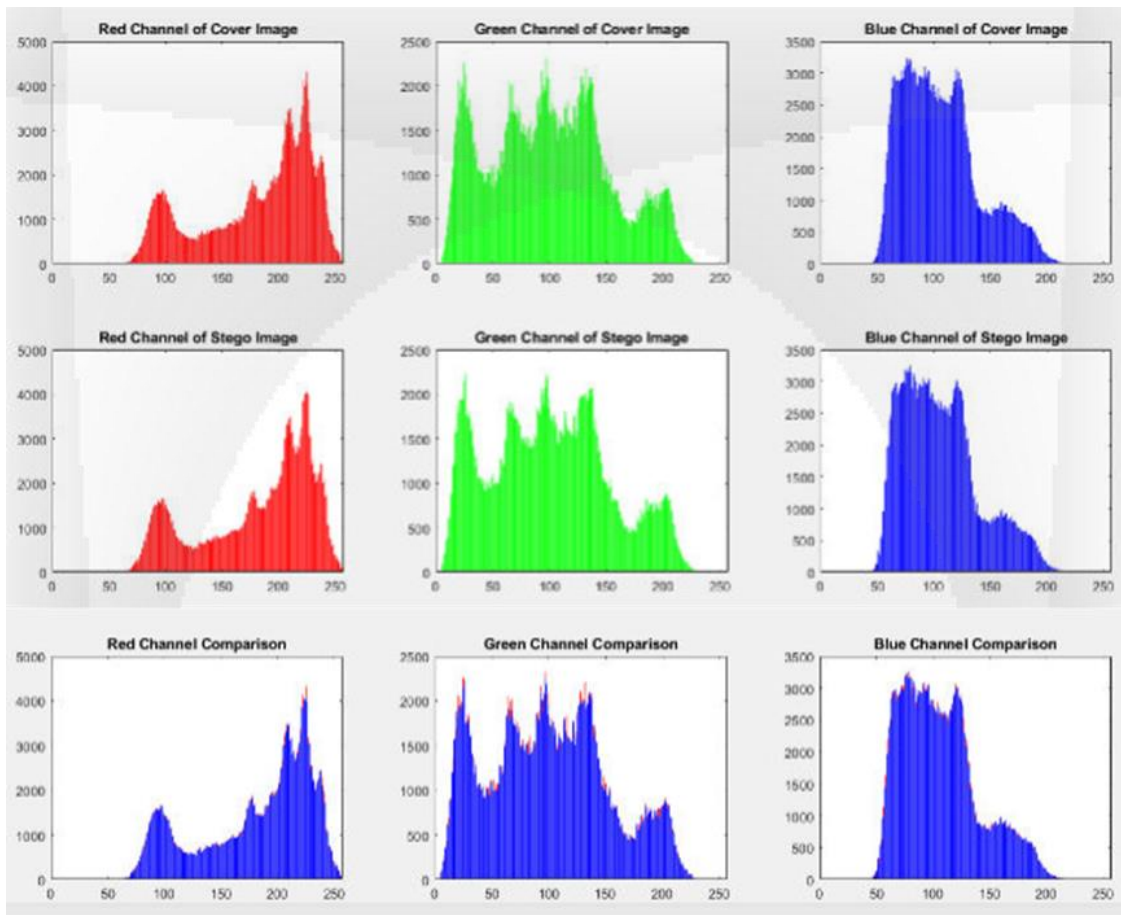


Fig. 10. Three-colour channel histogram analysis in Lena set

The overall shape and value of the histograms between the cover image and the stego image were identical. Therefore, the difference between the histograms of the Lena cover image and the Lena stego image can be concluded as a tiny difference until no apparent changes can be detected in the stego image. Figure 11 shows the histogram difference between the cover image and the stego image in the Lena Set.

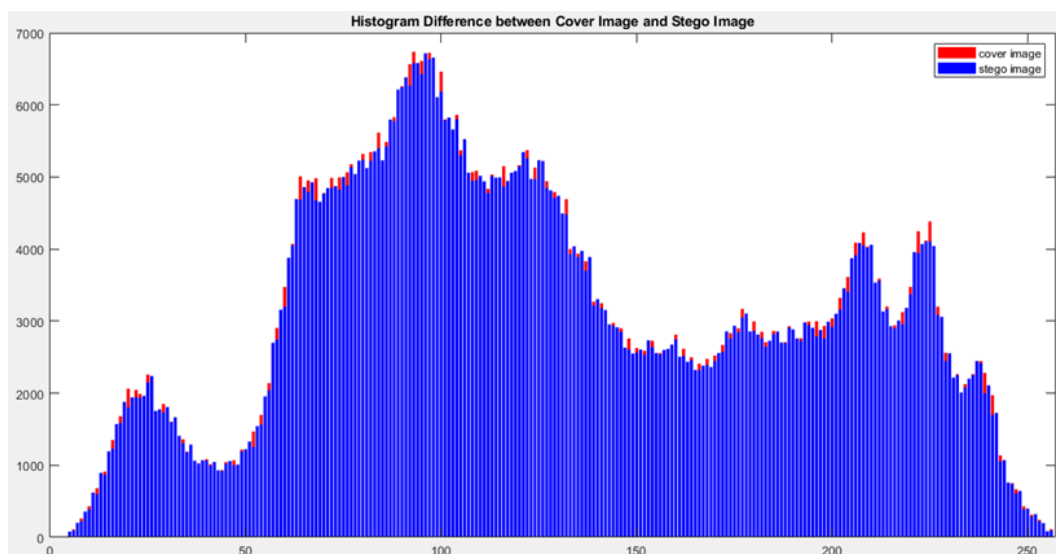


Fig. 11. Histogram difference in Lena set

Figure 12 shows the three colour channels histogram, including the comparison of the cover image and stego image in the Lake set. The overall shape and value of the histograms between the cover image and the stego image were identical. Like the Lena set, the difference between the histograms of the Lake set was very tiny until no obvious changes can be detected on the stego image.

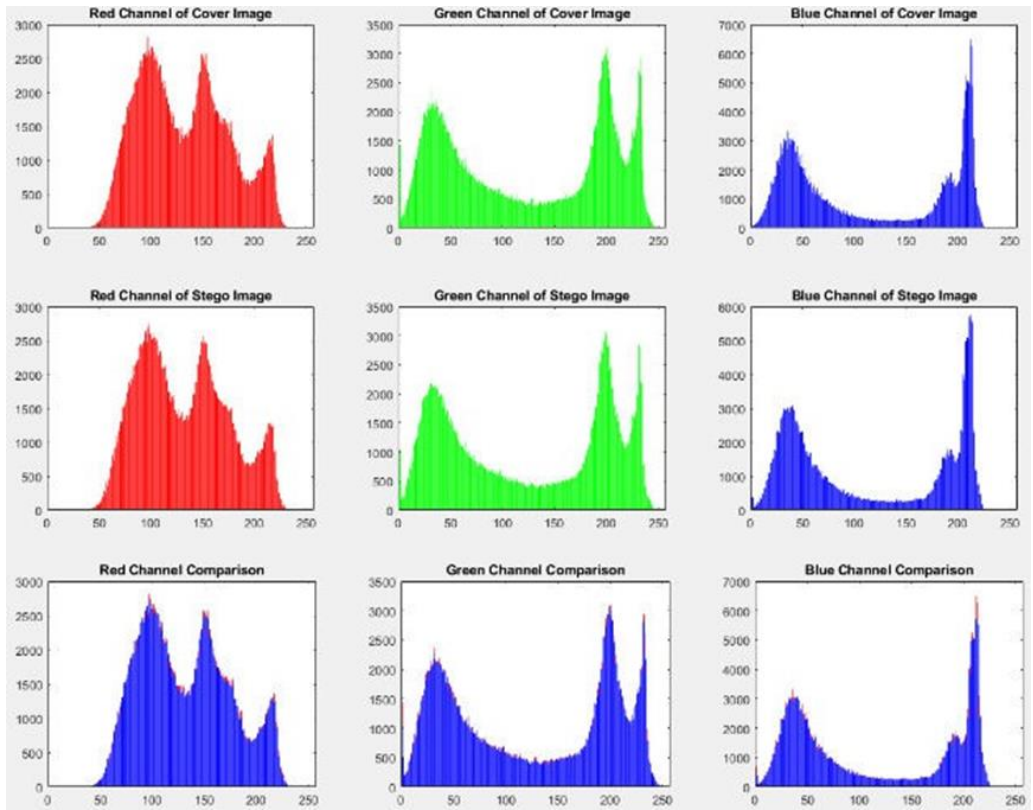


Fig. 12. Three-colour channel histogram analysis in Lake set

Figure 13 shows the histogram difference between the cover image and the stego image in the Lake set.

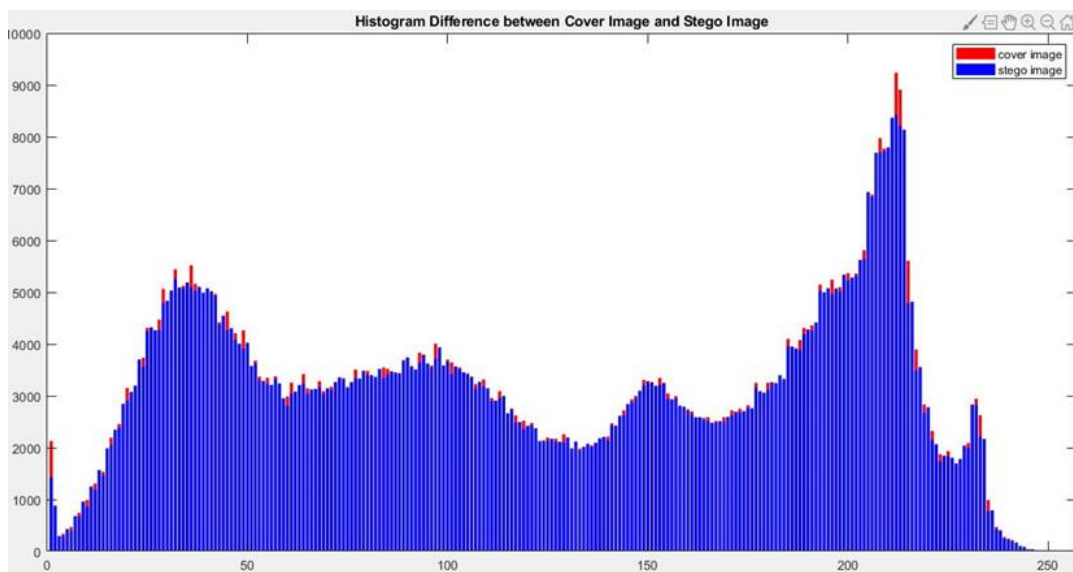


Fig. 13. Histogram difference in Lake set

4.3 Imperceptibility Evaluation

The encrypting size payload variable investigates the relationship between the imperceptibility and the payload capacity in the proposed image steganography method. In imperceptibility evaluation, the cover image: Lena and Lake was compared with a set of stego images having the payload capacity of 20 %, 40 %, 60 %, 80 % and 100 %. Figure 14 shows the Lena cover image with different encrypting payload sizes.



Fig. 14. Lena image set (a) Cover image (b) 20 % Stego image (c) 40 % Stego image (d) 60 % Stego image (e) 80 % Stego image and (f) 100 % Stego image

4.4 PSNR and SSIM Results

The Lena and Lake image sets of the cover image and stego images were evaluated by calculating the PSNR and the SSIM. The results were tabulated to investigate the imperceptibility of the proposed image steganography method. Table 1 shows the PSNR and SSIM of Lena image after embedding variable size payloads. Size of the ideal image was 750.42 kb. Actual size of 512 x 512 Lena image was 782.65 kb, and maximum embedding capacity for 512 x 512 image, 95.81 kb. The minimum acceptable value for PSNR was 30 dB to ensure that the overall visual structure in the digital image was not damaged. Table 1 shows that the PSNR values of the Lena image set were above the minimum acceptable value. The stego image with 100 % of payload capacity has the PSNR value of 51.19 dB.

The results of the Lake image set in Table 2 were also above the minimum acceptable value of PSNR. The Lake stego image with 100 % of payload capacity had a value of 51.15 dB in PSNR. Both PSNR results expressed that the proposed image steganography method performs well in the imperceptibility evaluation. Size of ideal 512 x 512 image was 750.42 kb. Actual size of 512 x 512 lake image was 783.13 kb, while maximum embedding capacity for 512 x 512 image, 95.81 kb.

Table 1

The PSNR and SSIM of Lena image after embedding variable size payloads

Payload (kb)	Payload capacity (%)	Image embedding capacity (%)	Proposed LSB steganography method	
			PSNR (dB)	SSIM (%)
19.16	20	2.45	58.11	99.96
38.33	40	4.90	55.10	99.92
57.49	60	7.35	53.36	99.88
76.65	80	9.79	52.14	99.84
95.80	100	12.24	51.19	99.81

Table 2

The PSNR and SSIM of Lake image after embedding variable size payloads

Payload (kb)	Payload capacity (%)	Image embedding capacity (%)	Proposed LSB steganography method	
			PSNR (dB)	SSIM (%)
19.16	20	2.45	57.73	99.97
38.33	40	4.89	54.92	99.95
57.49	60	7.34	53.23	99.92
76.65	80	9.79	52.04	99.89
95.80	100	12.23	51.12	99.86

The SSIM is the parameter that measures the similarity of the cover image and stego image. The higher the value of SSIM approaching 100 %, the higher the visual structure both images are identical. Table 1 shows the results of the Lena image set in SSIM having a value of more than 99.8 %, so as the results of the Lake image set in Table 2. These results can conclude that the differences between the cover image and the stego image in the proposed image steganography method did not exceed 0.2 %.

From Figure 16 and 17, the imperceptibility of the proposed image steganography method was inversely proportional to the percentage of payload capacity encrypted in the stego image. However, the minimum values of the imperceptibility parameter of the proposed image steganography were classified as an imperceptible method in steganography. From the relationship between the imperceptibility and payload capacity, the imperceptibility of the stego image can be reduced using a bigger size of the cover image to increase the size of payload capacity. The same secret message encrypted in the cover image will be taking a smaller percentage of payload capacity, thus increasing the PSNR and SSIM value.

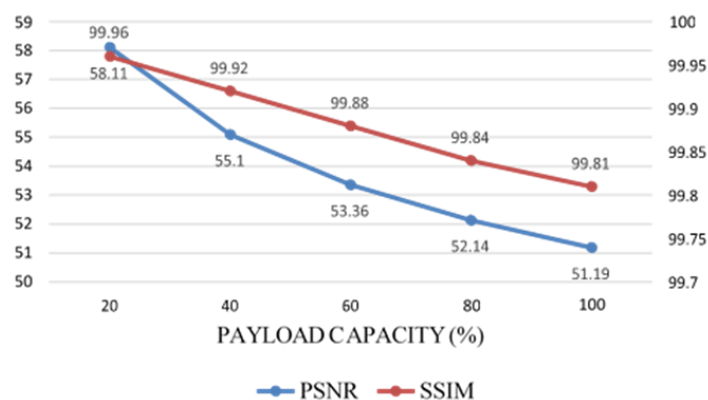


Fig. 16. Graph of Lena image set's PSNR and SSIM evaluation

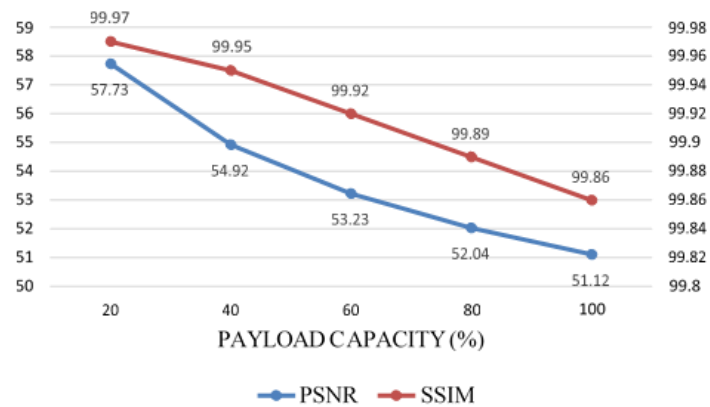


Fig. 17. Graph of Lake image set's PSNR and SSIM evaluation

5. Conclusion

In conclusion, this research has proposed an image steganography method implementing the LSB substitution to conceal and detect secret messages in a digital colour image. The proposed image steganography will then use a lossless .png image as the cover image and accepts multiple files as the secret message during encryption. The LSB substitution will replace every LSB of data in the cover image with the secret message in binary format.

The LSB substitution is a good technique in image steganography with high imperceptibility. However, the ordinary LSB substitution method is vulnerable to steganalysis. Combining the symmetric key cryptography with the LSB substitution has overcome the weaknesses. SHA-256 and AES-128 with the secret key will protect the confidentiality of the secret message encrypted. The random substitution referring to the secret key will decrease the chances of the stego image being suspected.

At the end of the research, the proposed image steganography achieved the research objectives: secure and safe image steganography using the LSB substitution method, which can accept multiple different types of secret messages to generate an imperceptible stego image with secret key protection using AES-128.

References

- [1] Gupta, Gaurav, Shruti Patil, Piyush Varma, Ankita Raikwade, and Yogita Sawant. "Information Security Through Image Processing by Utilizing Steganography."
- [2] Kumar, Arvind, and Km Pooja. "Steganography-A data hiding technique." *International Journal of Computer Applications* 9, no. 7 (2010): 19-23. <https://doi.org/10.5120/1398-1887>
- [3] Liddell H.G., Scott R., Jones H.S., McKenzie R.. (1984). A Greek-English Lexicon. Oxford University Press.
- [4] Verma, Vidya, Abhishek Kajal, and Isha Kajal. "Enhancement of payload capacity for image steganography based on LSB." *Int. J. Comput. Appl. Technol. Res* 5, no. 10 (2016): 678-682.
- [5] Rout, Haripriya, and Brojo Kishore Mishra. "Pros and cons of cryptography, steganography and perturbation techniques." *IOSR Journal of Electronics and Communication Engineering* 76 (2014): 81.
- [6] Sarkar, A., and S. Karforma. "Image steganography using password based encryption technique to secure e-banking data." *International Journal of Applied Engineering Research* 13, no. 22 (2018): 15477-15483.
- [7] Kekre, H. B., Archana Athawale, and Pallavi N. Halarnkar. "Increased capacity of information hiding in LSBs method for text and image." *International Journal of Electrical, Computer and Systems Engineering* 2, no. 4 (2008): 246-249.
- [8] Rout, Haripriya, and Brojo Kishore Mishra. "Pros and cons of cryptography, steganography and perturbation techniques." *IOSR Journal of Electronics and Communication Engineering* 76 (2014): 81.
- [9] Amarendra K, V. N. Mandhala, B. C. Gupta, G. G. Sudheshna, V. V. Anusha. (2019). Image Steganography Using LSB. *International Journal of Scientific & Technology Research* vol. 8, Issue 12, December 2019, ISSN 2277-8616
- [10] Bansod, Sneha, and Gunjan Bhure. "Data encryption by image steganography." *Int. J. Inform. Comput. Technol. Int. Res. Publ. House* 4 (2014): 453-458.

- [11] Kaur, Harpreet, and Jyoti Rani. "A Survey on different techniques of steganography." In *MATEC web of conferences*, vol. 57, p. 02003. EDP Sciences, 2016. <https://doi.org/10.1051/mateconf/20165702003>
- [12] Rout, Haripriya, and Brojo Kishore Mishra. "Pros and cons of cryptography, steganography and perturbation techniques." *IOSR Journal of Electronics and Communication Engineering* 76 (2014): 81.
- [13] Singh, Arun Kumar, Juhi Singh, and Harsh Vikram Singh. "Steganography in images using LSB technique." *International Journal of Latest Trends in Engineering and Technology (IJLTET)* 5, no. 1 (2015): 426-430.
- [14] Wang, Yaojie, Xiaoyuan Yang, and Hengkang Jin. "Generative Image Steganography Based on GANs." In *Trusted Computing and Information Security: 13th Chinese Conference, CTCIS 2019, Shanghai, China, October 24–27, 2019, Revised Selected Papers 13*, pp. 1-15. Springer Singapore, 2020. https://doi.org/10.1007/978-981-15-3418-8_1
- [15] Sumathi, C. P., T. Santanam, and G. Umamaheswari. "A study of various steganographic techniques used for information hiding." *arXiv preprint arXiv:1401.5561* (2014).