Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13187

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Optimizing Digital Image Steganography through Hybridization of LSB and Zstandard Compression

Fandy Novanto[1]*, Adhitya Nugraha[2], Joshua Calvin Kurniawan[3], Ariel Emmanuel Prayogo[4]
[1,2,3,4]Universitas Dian Nuswantoro, Indonesia
[1] fandy2832@gmail.com, [2] adhitya@dsn.dinus.ac.id, [3] kurniawanjoshua05@gmail.com,
[4]arielemmanuel4@gmail.com

Submitted : Nov 21, 2023 | **Accepted** : Nov 24, 2023 | **Published** : Jan 1, 2024

Abstract: In response to the growing need for secure digital communication globally, this research delves into an innovative strategy for enhancing data transmission security through steganography. This inventive approach involves the integration of the conventional Least Significant Bit (LSB) method with Zstandard (Zstd) compression to elevate the quality of stego images. The study carefully explores how the synergistic use of LSB and Zstd contributes to an improved equilibrium between embedding capacity and visual quality in stego images. This hybrid methodology capitalizes on the efficiency of Zstd in reducing file size, thereby facilitating more effective data concealment using LSB. The experimental outcomes showcase a notable 51.2% increase in embedding capacity, a 4.70% elevation in PSNR value, accompanied by a substantial 51.03% decrease in MSE value. Additionally, SSIM values hover around 0.007%, indicating a perceptually minimal difference between the original and steganographically modified images. These compelling results underscore the efficacy of the proposed method, highlighting its proficiency in preserving and enhancing the quality of stego images generated through the embedding process. This research signifies a significant stride in the realm of secure digital communication, demonstrating a promising fusion of traditional LSB with advanced Zstd compression for optimizing digital image steganographic.

Keywords: *Secure Digital Communication;* Steganography; LSB; Zstd; *Data Concealment*

INTRODUCTION

The information era has facilitated the widespread use of the internet and various forms of digital data, including text, images, videos, and audio, in everyday life. The significant amount of crucial data transferred over the internet every second emphasizes the importance of secure methods to protect data transmission. Vulnerabilities in communication networks pose risks, allowing unauthorized users to intercept or eavesdrop on transmitted data. However, these risks can be mitigated through various methods, such as encryption and data hiding.

Steganography, a practice that enables the transmission of hidden messages, files, or sensitive data without arousing suspicion, commonly employs the Least Significant Bit (LSB) as one of the most prevalent methods (Hussain et al., 2018). LSB operates by modifying the least significant bit of a pixel, making it simple and comprehensive. Nevertheless, LSB has its limitations, especially in terms of capacity, as this method can only embed 1 bit per pixel (Zakaria et al., 2018). Specifically, in RGB 512x512 images serving as cover images, the conventional LSB method can only embed 12.28 kilobytes of data. This limitation increases the risk of distortion in data by using every pixel in the image (Walia et al., 2018; Wang et al., 2020).

The correlation between the size of hidden data and the quality of the stego image is a critical factor in the field of steganography. As the size of hidden data increases, there is often a trade-off with the visual quality of the stego image. A larger amount of hidden data can cause distortions or artifacts visible in the image, compromising the integrity and overall quality of the picture(Nashat & Mamdouh, 2019). Striking a balance between the amount of hidden data and maintaining visual quality in the stego image is a challenge that requires careful consideration.

This paper aims to introduce a new LSB method by combining LSB with Zstandard for RGB cover images. This hybrid technique leverages the efficiency of Zstandard in reducing file size, which is then combined with LSB to hide information in the least significant bits of the image. The proposed method can produce stego images with significantly higher quality compared to conventional approaches. To validate its effectiveness, experiments are conducted on several commonly used RGB cover images such as "Lena," "Baboon," and "Pepper".

LITERATURE REVIEW

Over the years, significant efforts have been made to enhance the efficiency of the Least Significant Bit (LSB) method in steganography. Various approaches have been implemented and tested, focusing on integrating LSB with different algorithms to optimize data hiding (Al-Azzeh et al., 2019; Rustad et al., 2022). A robust LSB method not only effectively conceals information but also minimizes the impact on the cover image's quality, ensuring that the stego-image remains visually close to the original. The most commonly chosen algorithm is compression, as reducing data minimizes the space needed for the embedding process in the cover image.

Integrating compression algorithms with the Least Significant Bit (LSB) method is a common practice in digital data concealment (AbdelWahab et al., 2019; Jayapandiyan et al., 2020; Tayyeh & Al-Jumaili, 2022; Walker et al., 2023). This approach is widely adopted because it efficiently optimizes storage space while hiding information. Integrating compression algorithms into LSB offers a dual advantage, reducing file size and covertly concealing data in the least significant bits of an image or file. This technique is often used in various applications, ranging from secure communication to digital watermarking (S et al., 2021; Venkatesh et al., 2023).

Various compression methods can be applied to data, and the use of compression to aid data concealment in steganography is common (Ibrahim Almazaydeh et al., 2018; Sethi & Patel, 2019). Moreover, other studies also employ compression techniques in multilevel image steganography(Rahman et al., 2023). Some use the Discrete Cosine Transform (DCT) compression algorithm to shrink and hide secret messages successfully (AbdelWahab et al., 2019; Jayapandiyan et al., 2020). Both approaches have produced high-quality images and higher security levels compared to using LSB alone.

When comparing compression algorithms such as Zstandard with other compression algorithms like Huffman and LZ77, Zstandard can deliver superior performance (Ibrahim Almazaydeh et al., 2018; Rahman et al., 2023; Walker et al., 2023). Compared to Huffman and LZ77, Zstandard exhibits superior compression performance, combining speed and efficiency to achieve smaller file sizes without sacrificing processing speed. Zstandard implements the best string selection algorithm and dynamic programming, resulting in smaller file sizes with optimal processing speed. This makes Zstandard a preferred choice in scenarios where compression ratio and speed are crucial considerations (Walker et al., 2023).

METHOD

This research aims to introduce an update to the Least Significant Bit (LSB) method through hybridization with the Zstandard compression algorithm. The experiment comprises two stages: the embedding process, where data is embedded into the image, and the extraction process, where the embedded data can be retrieved. The concealed data in this experiment is alphanumeric text data, similar to the type commonly used in everyday life.

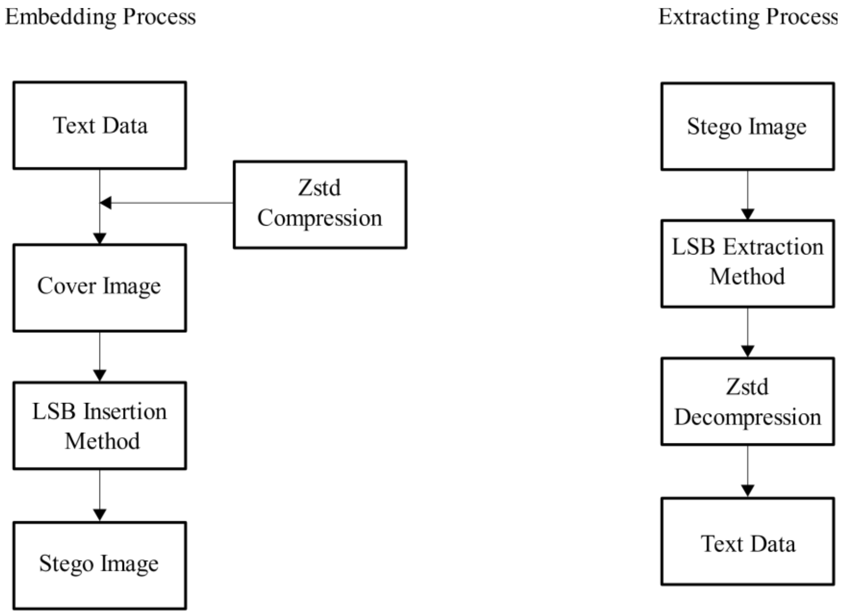

Figure 1. Embedding and extracting process

Dataset

The data used for the embedding process is randomly generated alphanumeric text data with sizes ranging from 1000 to 5000 bytes. The text data is created by extracting random text segments to simulate the writing style commonly done by humans in literature. This text data undergoes compression using the Zstandard algorithm to reduce its size, minimizing the number of pixels used during the embedding process in the cover image. Three images, each sized 512 x 512 pixels, namely "Lena," "Baboon," and "Pepper," will be used as cover images in this experiment.

Zstandard Algorithm

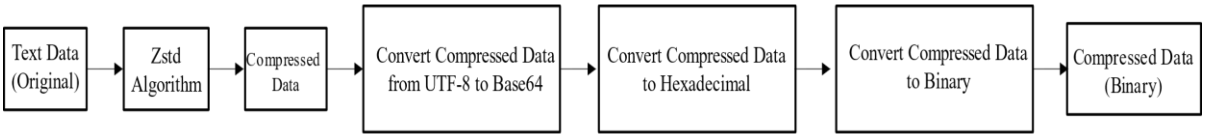
Text Data (Original)
Zstd Algorithm
Compressed Data
Convert Compressed Data from UTF-8 to Base64
Convert Compressed Data to Hexadecimal
Convert Compressed Data to Binary
Compressed Data (Binary)

Figure 2. Zstd compression process

The Zstandard algorithm (Zstd) is a compression algorithm that combines entropy encoding using the Finite State Entropy (FSE) algorithm and Huffman Encoding. Firstly, Zstd encodes text data using the FSE algorithm, and the encoded result is then compressed using Huffman Encoding. By using this combination, Zstd produces efficient compression, resulting in significantly smaller text data sizes. The Zstd compression algorithm is employed to compress text data before using it in the embedding process.

LSB

Least Significant Bit (LSB) is a steganographic technique that involves replacing the least significant bit of pixel values in an image with secret information. In the context of digital images, each pixel is represented by a combination of red, green, and blue color channels. By modifying the least significant bit, which has the least impact on the pixel's overall color, LSB embeds hidden data without noticeably altering the visual appearance of the image. The process is reversible, allowing the extraction of the concealed information by retrieving the modified least significant bits from the steganographed image.

Embedding Process

The embedding process is conducted using the LSB method, replacing the least significant bit of the binary value in RGB pixels. Before undergoing the embedding process, the text data to be embedded undergoes compression using the Zstd algorithm to reduce the text data size. Subsequently, the data is converted into binary form, and each bit is embedded into each color component of the RGB pixel of the cover image. The embedding process will only stop when all bits in the binary row are exhausted.

Stego-Image

After the embedding process, a stego-image is produced, incorporating the concealed text data within the original image. To evaluate the quality of the stego-image, metrics such as Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) are employed. These metrics provide insights into the level of distortion introduced during the embedding, helping to quantify the trade-off between data hiding capacity and the visual fidelity of the stego-image. Assessing these metrics allows for a comprehensive analysis of the effectiveness of the steganographic technique employed.

$$MSE = (1/n)\ \Sigma\ (Y_i - \bar{Y}_i)^2 (1)$$

The Mean Squared Error (MSE) measures the average squared difference between corresponding pixels of the original and steganographed images. In the formula(1), each pixel's difference is squared, and the resulting values are averaged to quantify the overall distortion. A lower MSE indicates a closer resemblance between the original and steganographed images, while a higher MSE implies greater dissimilarity.An MSE value approaching 0 dB indicates minimal deviation from the original image used in the embedding process.

$$PSNR = 10\log_{10}((MAX_i^2) / MSE) \qquad (2)$$

Formula(2) is a quantitative measure used to assess the quality of a stego-image compared to the original image. It calculates the ratio of the maximum possible pixel value squared to the Mean Squared Error (MSE), representing the power of the original image to the power of the noise introduced during embedding. The logarithmic scale and multiplication by 10 make the PSNR a convenient metric, providing a higher value for less perceptual distortion and better image quality preservation in the stego-image.. PSNR values in the range of 30 dB to 50 dB indicate very high stego-image quality.

$$SSIM = (2 * \mu x * \mu y + c_1) * (2 * \sigma xy + c_2) / (\mu x^2 + \mu y^2 + c_1) * (\sigma x^2 + \sigma y^2 + c_2) \qquad (3)$$

The Structural Similarity Index (SSIM) is a metric used to assess the similarity between two images, evaluating their luminance, contrast, and structure. The formula incorporates means (μx and μy), standard deviations (σx and ay), and cross-covariance (σxy) of the original and stego-images. By considering these factors, SSIM provides a comprehensive measure of perceptual similarity, offering insights into how well the stego-image preserves the visual characteristics of the original. Furthermore, an SSIM value approaching 1 or 100% indicates that the stego-image is nearly identical or very similar to the original image (This journal uses 1 as an indicator of identical similarity).

RESULT

Two experiments were conducted to compare the performance of the proposed method: one using the conventional LSB method and the other using the proposed method. The experiments utilized images of "Lena," "Baboon," and "Pepper," with text data ranging from 1000 bytes to 5000 bytes being embedded. The experimental results are presented in tables, with a comparative analysis between the common LSB method and LSB with Zstd compression. Tables 1 and 2 illustrate how Zstd compression effectively reduces the size of data, thereby enhancing the image capacity for message concealment. Another table provides a comparison between the regular LSB method and LSB with Zstd compression, using PSNR, MSE, and SSIM values as metrics to assess stego-image quality.

Table 1 .Compression Result using Zlib Compression

Size Before Compression (Byte)	Size After Compression (Byte)	Compression Rate
1000	560	43,99%
2000	1003	49,85%
3000	1437	52,09%
4000	1855	53,61%
5000	2176	56,48%

Table 1 displays the sizes of various text data, ranging from 1000 bytes to 5000 bytes, before and after compression using the Zstd algorithm. Additionally, the table provides the compression ratio in percentage form. The compression ratio signifies that the Zstd algorithm is capable of reducing the data size by the specified percentage.

$$\text{Compression Rate (\%)} = ((x - y) / x) * 100\% \qquad (4)$$

The compression ratio formula in percentage represents the reduction in data size by subtracting the compressed size (y) from the original size (x), dividing the result by the original size, and multiplying by 100 to express it as a percentage. From the table above, it can be observed that there is a minimum 40% reduction in size for each data to be embedded into the cover image.

Table 2. Comparison of Pixel Usage Between the Common LSB Method and the LSB Compression Using Zstd Result

Hidden Data(Byte)	Convensional LSB Pixel Usage	LSB+Zstd Pixel Usage
1000	2667	1494
2000	5334	2675
3000	8000	3832
4000	10667	4947
5000	13334	5803

Table 2 illustrates the difference in pixel utilization between the regular LSB method and the LSB method using Zstd compression.

$$\text{Pixel Usage}=(x*8)/3 \qquad (5)$$

The formula for calculating pixel usage in LSB (Least Significant Bit) is obtained by multiplying the total byte of hidden data(x) by 8 to convert it into bits and then dividing it by the color components in each pixel, which in this case is 3, as it is an RGB image. It can be observed that the proposed method effectively reduces the usage of pixels available in the cover image and this reduced pixel usage enhances the quality and integrity of the message embedded in the stego-image. The proposed method will yield superior stego-image quality because the changes in pixel values in the cover image will also be minimal, minimizing the likelihood of artifacts in the stego-image.

Table 3. Comparison Between Conventional LSB Method and LSB Compression Using Zstd Results

		Conventional LSB	LSB +ZSTD

Cover Image	Data (Byte)	PSNR (dB)	MSE (dB)	SSIM	PSNR (dB)	MSE (dB)	SSIM
Lena	1000	71,03544	0,00512	0,9999786	73,49665	0,00291	0,9999924
	2000	68,01438	0,01027	0,9999332	71,03652	0,00512	0,9999790
	3000	66,25203	0,01541	0,9998856	69,44192	0,00739	0,9999597
	4000	65,03889	0,02038	0,9998406	68,36294	0,00948	0,9999390
	5000	63,32639	0,03023	0,9997565	66,92001	0,01322	0,9999009
Baboon	1000	71,10389	0,00504	0,9999994	73,50807	0,00290	0,9999998
	2000	68,13098	0,01000	0,9999979	71,04299	0,00511	0,9999994
	3000	66,36786	0,01501	0,9999961	69,52714	0,00725	0,9999988
	4000	65,10827	0,02006	0,9999944	68,37519	0,00945	0,9999981
	5000	63,37840	0,02987	0,9999907	66,93549	0,01317	0,9999969
Pepper	1000	70,99040	0,00518	0,9999922	73,54633	0,00287	0,9999972
	2000	67,97797	0,01036	0,9999782	70,98827	0,00518	0,9999923
	3000	66,22702	0,01550	0,9999649	69,53782	0,00723	0,9999863
	4000	65,00300	0,02055	0,9999490	68,34260	0,00952	0,9999799
	5000	63,29182	0,03047	0,9999139	66,91499	0,01323	0,9999696

Table 3 presents an analysis of image quality comparison between the conventional LSB method and the LSB method with Zstd compression, both applied to the cover images "Lena," "Baboon," and "Pepper." The results show a significant improvement in image quality. From the data presented in Table 3, the LSB method with Zstd compression can significantly enhance the quality of the stego-image. As shown in the table, the proposed method successfully increases the PSNR value of the stego-image significantly.

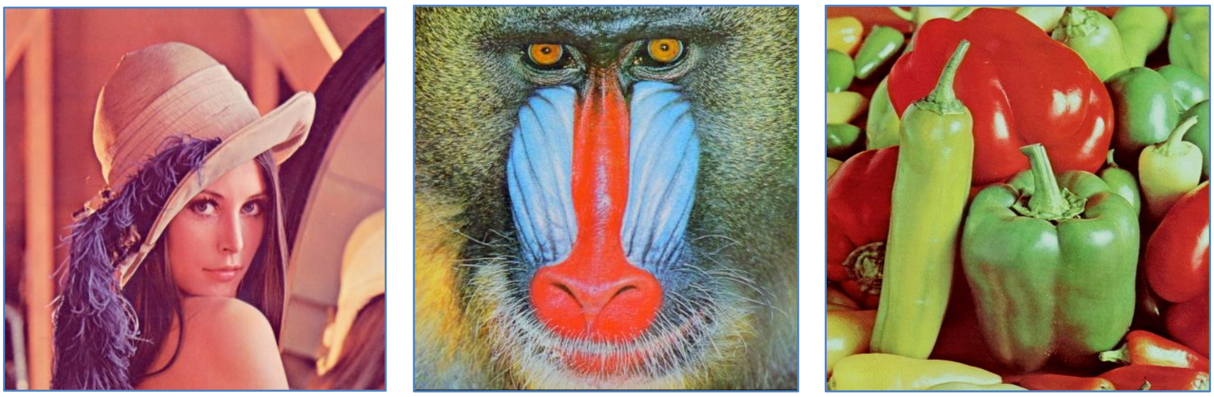

Figure 3. Original image used for the embedding process

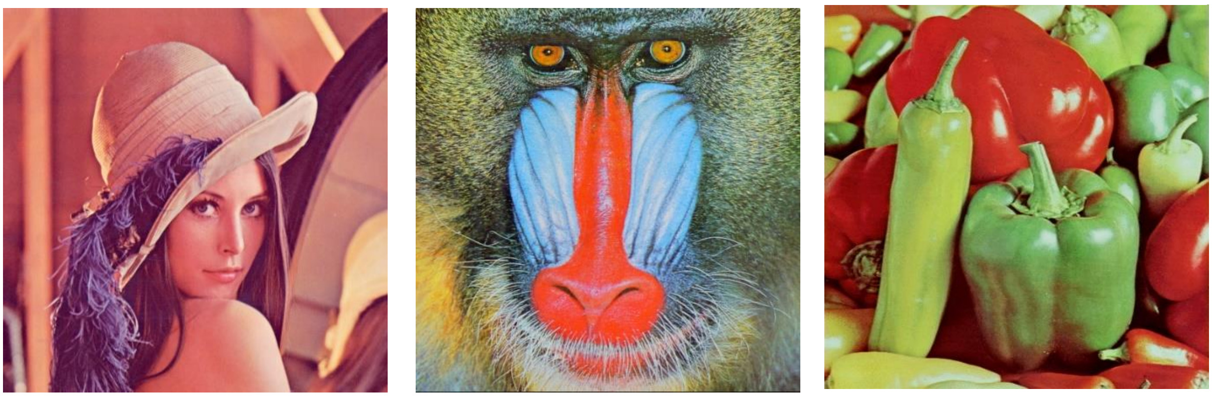

Figure 4. Image after the embedding process

Figure 3 and 4 shows the difference between the original image and the image after the embedding process is done. The image shown in both figure shows that both the original image and the image after the embedding process is identical and undetectable with human eyes, which confirm the SSIM value in table 3 that present the SSIM value of the stego-image is identical to the original image used for the embedding process

DISCUSSIONS

On the "Lena" cover image, the results show a significant improvement in image quality. From the data presented in Table 3, the LSB method with Zstd compression can significantly enhance the quality of the stego-image, as evident in the table. The proposed method successfully increases the PSNR value of the stego-image significantly. For example, using 1000 bytes of data, the regular LSB method only has a PSNR score of 71.03544 dB, while the proposed method achieves a notable increase with a PSNR score of 73.49665 dB. Other measurement metrics also indicate a significant improvement, with a decrease in MSE values and an increase in SSIM values across all data usage sizes.

Using the "Pepper" image as the cover image exhibits a remarkable improvement in image quality achieved with the LSB method with Zstd compression compared to the conventional LSB method. From the presented data, the proposed method achieves a maximum MSE value of 0.03047 dB, which is still considered impressive compared to the conventional LSB method, which, with a data size of 5000 bytes, reaches 0.01323 dB.

Similarly, using the "Baboon" image as the cover image demonstrates a significant improvement in image quality with the LSB method with Zstd compression compared to the traditional LSB method. The SSIM metric serves as a tool to measure the perceptual differences between the original Pepper image and the steganographically modified version, providing insights into the effectiveness and visual impact of LSB-based data hiding techniques. According to the data, the proposed method achieves the highest SSIM value of 0.9999998, indicating a significant improvement compared to the regular LSB method, which, with the same data size of 1000 bytes, achieves a value of 0.9999994.

The utilization of the LSB method alongside the Zstd compression algorithm has proven its effectiveness in expanding the cover image's capacity to accommodate additional embeddable data while simultaneously enhancing the quality of the stego-image. By employing Zstd compression, the size of text data is reduced, resulting in a substantial decrease in the pixels required for embedding, thereby creating much more space within the cover image. Consequently, the quality of the stego-image improves significantly as the number of pixels affected by the embedding process becomes minimal.

Using PSNR, MSE, and SSIM as measurement metrics, the experiments demonstrate impressive scores for the proposed method with all three images. With 5000 text data, the proposed method achieves a minimum PSNR score of 66.91499 dB using the "Pepper" image. From the perspective of MSE, the proposed method successfully maintains values close to 0 dB, with an MSE of 0.01323 dB using the "Pepper" image, while the other two images both have nearly identical scores, which is considered highly impressive. Furthermore, the proposed method consistently retains SSIM values around 0.9999 points, indicating a high level of similarity.

CONCLUSION

Based on the findings from the integration of the Least Significant Bit (LSB) method with Zstandard (Zstd) compression, it is evident that the combined approach is effective in enhancing steganographic applications in digital images. The experiments demonstrated an I ncrease in embedding capacity and an improvement in visual quality, as indicated by a 4.70% increase in PSNR value and a 51.03% decrease in MSE value. Additionally, the SSIM values experienced a 0.007% increase and even fell within the range of nearly identical. These results indicate that the proposed method successfully enhances the quality of the stego-image. Therefore, it can be concluded that the combination of LSB with Zstd compression not only improves the capacity of cover images but also enhances the quality of the resulting stego-image.

REFERENCES

AbdelWahab, O. F., Hussein, A. I., Hamed, H. F. A., Kelash, H. M., Khalaf, A. A. M., & Ali, H. M. (2019). Hiding data in images using steganography techniques with compression algorithms. *Telkomnika (Telecommunication Computing Electronics and Control)*, *17*(3), 1168–1175. https://doi.org/10.12928/TELKOMNIKA.V17I3.12230

Al-Azzeh, J., Alqadi, Z., Ayyoub, B., & Sharadqh, A. (2019). *Improving the Security of LSB Image Steganography*.

Hussain, M., Wahab, A. W. A., Idris, Y. I. Bin, Ho, A. T. S., & Jung, K. H. (2018). Image steganography in spatial domain: A survey. *Signal Processing: Image Communication*, *65*, 46–66. https://doi.org/10.1016/j.image.2018.03.012

Ibrahim Almazaydeh, el A., Sheshadri, H. S., & Padma, S. (2018). A Novel Method to Hide a Text in an Image Using a Dynamic Symmetric Key and Huffman Coding. *IJRECE*, *6*(3), 1095–1101.

Jayapandiyan, J. R., Kavitha, C., & Sakthivel, K. (2020). Optimal Secret Text Compression Technique for Steganographic Encoding by Dynamic Ranking Algorithm. *Journal of Physics: Conference Series*, *1427*(1). https://doi.org/10.1088/1742-6596/1427/1/012005

Nashat, D., & Mamdouh, L. (2019). An efficient steganographic technique for hiding data. *Journal of the Egyptian Mathematical Society*, *27*(1). https://doi.org/10.1186/s42787-019-0061-6

Rahman, S., Uddin, J., Hussain, H., Ahmed, A., Khan, A. A., Zakarya, M., Rahman, A., & Haleem, M. (2023). A Huffman code LSB based image steganography technique using multi-level encryption and achromatic component of an image. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-41303-1

Rustad, S., Setiadi, D. R. I. M., Syukur, A., & Andono, P. N. (2022). Inverted LSB image steganography using adaptive pattern to improve imperceptibility. *Journal of King Saud University - Computer and Information Sciences*, *34*(6), 3559–3568. https://doi.org/10.1016/j.jksuci.2020.12.017

S, S., Karanth, S., & Kumar, S. (2021). Protection of data using image watermarking technique. *Global Transitions Proceedings*, *2*(2), 386–391. https://doi.org/10.1016/j.gltp.2021.08.035

Sethi, N., & Patel, P. (2019). Steganography Technique with Huffman Code. *International Journal of Recent Technology and Engineering*, *8*(2S4), 867–870. https://doi.org/10.35940/ijrte.b1173.0782s419

Tayyeh, H. K., & Al-Jumaili, A. S. A. (2022). A combination of least significant bit and deflate compression for image steganography. *International Journal of Electrical and Computer Engineering*, *12*(1), 358–364. https://doi.org/10.11591/ijece.v12i1.pp358-364

Venkatesh, M. K., Sailaja, G., Tapaswi, C., Chandramani, G., & Rohith, C. (2023). *DIGITAL WATERMARKING TECHNIQUES USING LSB*.

Walia, G. S., Makhija, S., Singh, K., & Sharma, K. (2018). Robust stego-key directed LSB substitution scheme based upon cuckoo search and chaotic map. *Optik*, *170*, 106–124. https://doi.org/10.1016/j.ijleo.2018.04.135

Walker, L. A., Li, Y., Mcglothlin, M., & Cai, D. (2023). A Comparison of Lossless Compression Methods in Microscopy Data Storage Applications. *Proceedings of The 6th International Conference on Software Engineering and Information Management. (ICSIM '23)*, *14*(7), 129–137. https://doi.org/10.1101/2023.01.24.525380

Wang, Y., Tang, M., & Wang, Z. (2020). High-capacity adaptive steganography based on LSB and Hamming code. *Optik*, *213*. https://doi.org/10.1016/j.ijleo.2020.164685

Zakaria, A. A., Hussain, M., Wahab, A. W. A., Idris, M. Y. I., Abdullah, N. A., & Jung, K. H. (2018). High-capacity image steganography with minimum modified bits based on data mapping and LSB substitution. *Applied Sciences (Switzerland)*, *10*(11). https://doi.org/10.3390/app8112199