Imperceptibility and Payload Capacity Analysis of LSB and DCT Steganography on Digital Images

Mahesa Satria Prayata - 18223082
Program Studi Sistem dan Teknologi Informasi
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jalan Ganesha 10 Bandung
E-mail: mahesaprayata@gmail.com , 18223082@std.stei.itb.ac.id

***Abstract*—Steganography hides secret information inside ordinary cover media so that the very existence of the message is concealed. For digital images, the two most widely used embedding domains are the spatial domain, represented by Least Significant Bit (LSB) substitution, and the frequency domain, represented by Discrete Cosine Transform (DCT) coefficient embedding. This paper presents a controlled experimental comparison of the two techniques along two axes that govern their practical usefulness: imperceptibility, measured by Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM), and payload capacity, measured in bits and in bits per pixel (bpp). Both methods were implemented from scratch in Python and evaluated on eight standard grayscale test images at five embedding levels each (10%, 25%, 50%, 75%, and 100% of capacity). The results show that LSB offers roughly 64 times the capacity of the block-based DCT scheme while also producing higher PSNR at every operating point, whereas DCT distributes more distortion per embedded bit because each embedded bit perturbs an entire 8×8 block. Both techniques remain visually imperceptible (PSNR above 44 dB, SSIM above 0.97) across all tested levels. The findings quantify the capacity–imperceptibility trade-off and provide a reproducible baseline for the companion steganalysis study**

Keywords—steganography; LSB; DCT; imperceptibility; payload capacity; PSNR; SSIM

I. Introduction

The growth of digital communication has made the confidential exchange of information a central concern of information security. Cryptography protects the content of a message by rendering it unintelligible, but the presence of ciphertext can itself attract attention and invite attack. Steganography takes a complementary approach: it conceals the existence of the message by embedding it inside an innocuous carrier, called the cover, producing a stego object that is perceptually indistinguishable from the original. Digital images are especially attractive cover media because they contain a large amount of perceptually redundant data that can be modified without producing visible artifacts.

Beyond covert person-to-person communication, image steganography underpins a range of both protective and adversarial applications. On the protective side it supports digital watermarking for copyright assertion, tamper-evident medical and military imaging, and the discreet transport of authentication tokens or integrity checksums. On the adversarial side, the same techniques are exploited to exfiltrate sensitive data past network monitors and to conceal command-and-control traffic inside ordinary-looking images. Understanding the imperceptibility and capacity limits of each embedding method therefore matters to defenders—who must know what an attacker can hide and how—as much as to legitimate users. The two techniques examined in this paper, LSB and DCT, are the archetypes of the two embedding domains and form the basis of most practical image-steganography systems in use today.

Two embedding strategies dominate image steganography. The first operates directly on pixel values in the spatial domain, its canonical form is Least Significant Bit (LSB) substitution, which replaces the lowest-order bit of selected pixels with message bits. The second operates in a transform domain, most commonly the Discrete Cosine Transform (DCT), embedding bits by modifying frequency coefficients of image blocks. The two approaches embody a fundamental design tension. Spatial embedding is simple and offers very high capacity but alters the image in ways that statistical analysis can later detect. Transform-domain embedding spreads each change across many pixels and is generally more robust, but it carries far less data and can introduce more distortion per embedded bit.

For a practitioner choosing between these techniques, two questions are decisive. First, how imperceptible is the stego image, that is, how closely does it resemble the cover under objective image-quality metrics? Second, how much secret data can each technique carry? These two properties are in direct conflict: embedding more data necessarily increases distortion. The purpose of this paper is to quantify that conflict experimentally.

The contribution of this work is a controlled, reproducible comparison of LSB and DCT steganography with respect to imperceptibility and payload capacity. Both algorithms were implemented independently, applied to a common set of cover images under identical conditions, and evaluated with three standard image-quality metrics (MSE, PSNR, SSIM) and two capacity measures (total bits and bits per pixel). The embedding load was swept across five levels for each method

so that the full capacity–quality curve could be observed rather than a single point.

The remainder of this paper is organized as follows. Section II reviews the theoretical background of image steganography and the evaluation metrics. Section III describes the design and implementation of the two embedding schemes and the experimental procedure. Section IV presents and analyzes the results. Section V concludes the paper.

II. Theoretical Background

A. Image Steganography

A steganographic system can be modeled as an embedding function that takes a cover image C, a secret message M, and optionally a key, and produces a stego image S that is perceptually equivalent to C but contains M. A corresponding extraction function recovers M from S. The quality of such a system is judged primarily by three properties: imperceptibility (the stego image must not reveal the presence of hidden data to a human observer or to statistical tests), capacity (the amount of data that can be embedded), and robustness (the ability of the message to survive processing of the stego image). This paper focuses on the first two properties.

B. Least Significant Bit (LSB) Substitution

LSB substitution is the simplest and most widely studied spatial-domain technique, and one of the earliest data-hiding methods used. Each pixel of a grayscale image is an 8-bit integer in the range [0, 255]. The least significant bit contributes a value of only 1 to the pixel intensity, so replacing it changes the pixel by at most one gray level, a change below the threshold of human visual perception. To embed a message, its bits are written sequentially into the least significant bits of successive pixels. Because exactly one bit can be stored per pixel, the raw capacity of an $H \times W$ grayscale image is $H \times W$ bits, the highest of any common scheme. The principal weakness of LSB substitution is that it disturbs the statistical regularities of the LSB plane in a way that dedicated steganalysis attacks can detect, which is not discussed in this paper.

C. Discrete Cosine Transform (DCT) Steganography

DCT-based steganography embeds information in the frequency domain rather than directly in pixel values. The image is divided into non-overlapping 8×8 blocks, and each block is transformed with the two-dimensional DCT which separates the block into a direct-current (DC) coefficient and a set of alternating-current (AC) coefficients representing increasing spatial frequencies. A message bit is then encoded by modifying a selected mid-frequency coefficient. Mid-frequency coefficients are chosen because the lowest frequencies carry most of the image energy (modifying them is highly visible) while the highest frequencies are fragile. In this work the bit is encoded by quantization index modulation (QIM): the chosen coefficient is quantized to the nearest multiple of a step Δ whose parity equals the message bit. Because each block stores only one bit, the capacity of the scheme is approximately $(H \times W) / 64$ bits, about one sixty-fourth that of LSB. After modification, the inverse DCT returns the block to the spatial domain.

D. Imperceptibility Metrics

Imperceptibility is quantified by comparing the stego image to the cover image with objective full-reference metrics.

The *Mean Squared Error* (MSE) is the average squared difference between corresponding pixels:

$$MSE = (1 / MN) \Sigma \Sigma [C(i, j) - S(i, j)]^2 \quad (1)$$

A smaller MSE indicates a stego image closer to the cover.

The *Peak Signal-to-Noise Ratio* (PSNR), expressed in decibels, relates the maximum possible pixel value to the embedding distortion:

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

where $MAX = 255$ for 8-bit images. Higher PSNR means greater imperceptibility; values above roughly 40 dB are generally considered visually indistinguishable from the original.

The *Structural Similarity Index* (SSIM) models the human visual system by comparing local luminance, contrast, and structure, returning a value in [0, 1] where 1 denotes a perfect match. SSIM often correlates better with perceived quality than PSNR because it is sensitive to structural distortion rather than to absolute pixel error alone.

E. Payload Capacity

Payload capacity measures how much secret data a technique can carry. It is reported here in two forms: the absolute number of embeddable message bits, and the *bits per pixel* (bpp), defined as the number of embedded bits divided by the total number of pixels. The bpp normalization allows images of different sizes to be compared on a common scale and exposes the intrinsic capacity difference between the two embedding domains.

F. Comparison of the Embedding Domains

The two domains differ in more than just where the bits are placed. Spatial-domain embedding such as LSB acts on individual, statistically independent pixels, which is what gives it both its very high capacity and its simplicity but it leaves a localized footprint in the least-significant-bit plane that statistical attacks can isolate. Transform-domain embedding such as DCT instead modifies coefficients that each summarize the joint behavior of 64 pixels. A single coefficient change is therefore diffused across the whole block when the image is reconstructed. This diffusion is the root cause of the trade-off measured in this paper, it makes each DCT bit costlier in distortion terms but also harder to localize, and it is the reason DCT-style embedding tends to survive operations such as mild compression that would destroy an LSB payload. The present study deliberately isolates the imperceptibility and capacity dimensions of this trade-off.

III. Design and Implementation

A. System Overview

The experimental system was implemented in Python 3 using NumPy for array processing, Pillow for image input/output, SciPy for the DCT, scikit-image for SSIM, and pandas and Matplotlib for tabulating and plotting the results. The codebase is organized into a reusable, side-effect-free library (`src/`) containing the embedding and metric routines, and a set of experiment scripts (`experiments/`) that perform all file input/output. All random operations are seeded so that every run is reproducible.

To ensure a fair comparison, both techniques operate on the grayscale representation of each cover image. This guarantees that LSB and DCT are evaluated on the same single-channel data with the same pixel count so the differences in results are attributable to the embedding method rather than to color handling.

B. LSB Embedding and Extraction

The LSB module embeds a payload by clearing the least significant bit of each pixel and replacing it with the next message bit, processing pixels in raster order. A 32-bit length header is prepended to the payload so that the extractor knows exactly how many bits to read, this makes extraction exact and avoids reliance on a fragile end-of-message delimiter. The usable capacity is therefore $H \times W - 32$ bits. Extraction reverses the process: it reads the 32-bit header to recover the payload length and then collects the corresponding number of least significant bits.

C. DCT Embedding and Extraction

The DCT module converts the cover to grayscale and processes it in non-overlapping 8×8 blocks. Each block is transformed with the orthonormal two-dimensional DCT. A single mid-frequency coefficient at block position (4, 3) is selected as the carrier. The message bit is embedded by quantization index modulation: the coefficient *c* is mapped to $q = round(c / \Delta)$ with quantization step $\Delta = 16$, and q is adjusted by one step if necessary so that its parity equals the message bit; the coefficient is then set to $q \times \Delta$. The inverse DCT returns the block to the spatial domain, and the result is clipped to [0, 255] and stored as an 8-bit image. The same 32-bit length header used by the LSB module frames the payload. The quantization step of 16 is large enough that the rounding incurred when the block is written back as 8-bit integers cannot flip the recovered bit, giving error-free extraction. The capacity is $(number\ of\ 8\times8\ blocks) - 32$ bits.

Two implementation parameters warrant justification. The carrier coefficient at position (4, 3) is a mid-frequency component, it sits far enough from the DC term that altering it does not visibly shift the block's overall brightness, yet it is not among the highest-frequency coefficients whose magnitudes are so small that quantization and rounding would render the embedded bit unreliable. The quantization step $\Delta = 16$ sets the trade-off between robustness and distortion. A larger step would make extraction even more reliable but would inject more distortion per bit. A smaller step would improve imperceptibility but risk bit errors once the block is rounded back to 8-bit integers. The value of 16 was chosen as the smallest step that yields error-free extraction across all test images so that the imperceptibility figures reported here are not inflated by a needlessly aggressive step.

D. Metric Computation

The metrics module computes MSE, PSNR, and SSIM directly from the definitions in Section II with PSNR returning infinity for identical images and SSIM computed with a data range of 255. Payload capacity is reported both as the total number of embedded bits and as bits per pixel.

E. Dataset and Experimental Procedure

Eight standard grayscale test images were used as covers, stored in the lossless TIFF format so that no lossy re-compression would affect the embedding or the subsequent quality measurements. The images span a range of sizes, from 256×256 to 1024×1024 pixels (approximately 0.066 to 1.05 megapixels), with the majority at 512×512. For each cover image and each technique, the payload was embedded at five levels corresponding to 10%, 25%, 50%, 75%, and 100% of that technique's capacity. The payload itself was a pseudo-random alphanumeric string generated from a fixed seed, ensuring identical payloads across runs. For each of the resulting 80 cases (8 images × 2 methods × 5 levels) the system saved the stego image and recorded the capacity in bits, the bits per pixel, the fill ratio, and the three imperceptibility metrics. The full per-case table was written to a CSV file, and PSNR-versus-bpp and SSIM-versus-bpp comparison plots were generated.

IV. Testing and Result Analysis

A. Imperceptibility Results

Table I reports the mean imperceptibility metrics over the eight cover images at each embedding level. The standard deviations were small on the order of ±0.05 dB for LSB PSNR and up to about ±0.55 dB for DCT PSNR, indicating that the results are consistent across images and not driven by outliers.

TABLE I. Mean Imperceptibility and Capacity by Method and Embedding Level

Method	Level	bpp	PSNR (dB)	SSIM	MSE
LSB	10%	0.1000	61.12	0.9997	0.0502
LSB	25%	0.2500	57.15	0.9992	0.1254
LSB	50%	0.4999	54.14	0,9984	0.2508
LSB	75%	0.7499	52.39	0.9978	0.3754
LSB	100%	0.9998	51.13	0.9971	0.5016
DCT	10%	0.0015	54.63	0.9987	0.2253
DCT	25%	0.0039	50.59	0.9963	0.5688
DCT	50%	0.0077	47.64	0.9931	1.1195

DCT	75%	0.0116	45.95	0.9905	1.6357
DCT	100$	0.0155	44.67	0.9876	2.2218

For both methods, PSNR and SSIM decrease monotonically and MSE increases as the embedding load grows, confirming the expected relationship between payload and distortion. Crucially, every configuration remains comfortably above the perceptual thresholds. The lowest mean values, reported in Table I, both occur for DCT at full capacity: 44.67 dB and SSIM 0.9876. Even the worst-performing individual image stayed high, across all eight images the minimum PSNR was 44.45 dB (the *stream_bridge* image at DCT full capacity) and the minimum SSIM was 0.980 (the *clock* image at DCT full capacity). In other words, neither technique produces visible artifacts at any tested embedding level, the differences between them are statistical and metric-based rather than perceptual.

Fig. 1 and Fig. 2 plot PSNR and SSIM against bits per pixel. Because the two methods operate over very different capacity ranges, their curves occupy disjoint regions of the horizontal axis (discussed further in Section IV-C).

Fig. 1. PSNR versus payload (bits per pixel) for LSB and DCT.

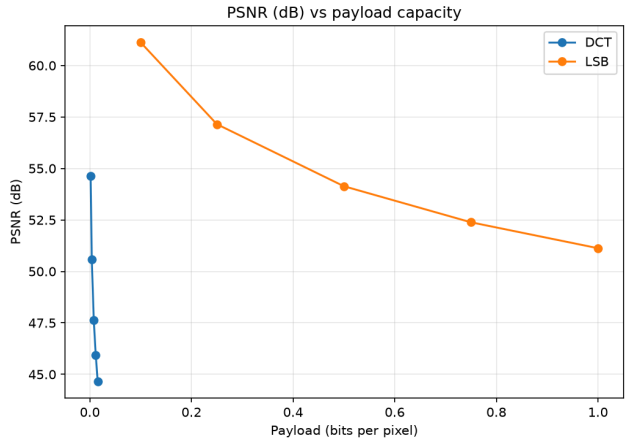
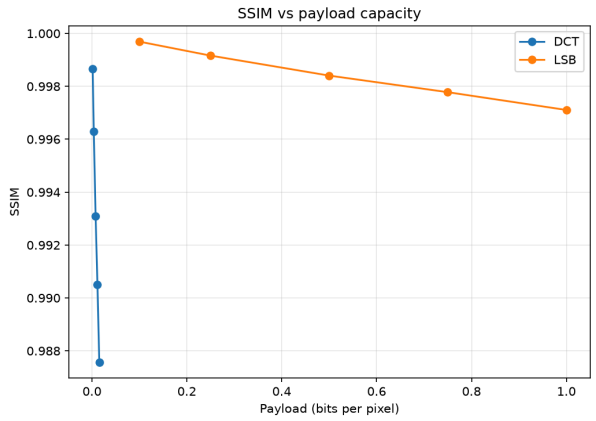
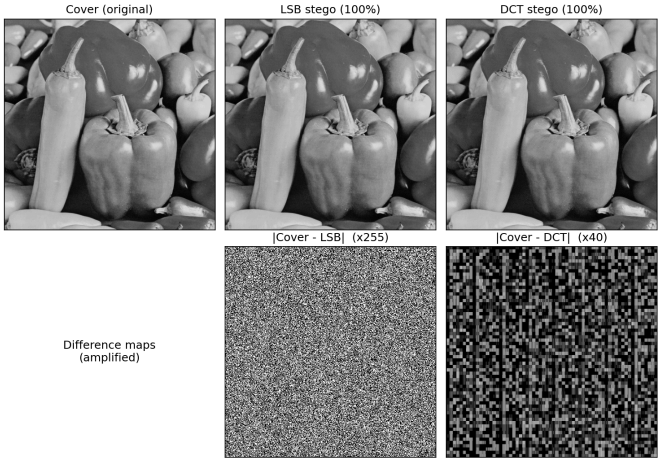


Fig. 2. SSIM versus payload (bits per pixel) for LSB and DCT.



A visual confirmation of imperceptibility is given in Fig. 3, which shows one cover image (the 512×512 *peppers* image) alongside its LSB and DCT stego versions at full capacity, together with the corresponding pixel-difference maps. The stego images are visually indistinguishable from the cover. The difference maps must be strongly amplified to be visible at all: the LSB difference is amplified by a factor of 255 (because every changed pixel differs by exactly one gray level) and the DCT difference by a factor of 40. The maps also reveal the structural distinction between the two methods. The LSB difference is a fine, pixel-level random pattern spread over the image, whereas the DCT difference exhibits a block structure aligned to the 8×8 grid, reflecting that each embedded bit perturbs a whole block. For this image the maximum absolute change of any pixel was 1 for LSB and 4 for DCT, consistent with the MSE figures in Table I.

Fig. 3. Visual comparison for the *peppers* cover. Top row: cover, LSB stego (100%), DCT stego (100%). Bottom row: amplified absolute-difference maps for LSB (×255) and DCT (×40).



B. *Payload Capacity Results*

The capacity difference between the two methods is large and structural. Averaged over the eight images, LSB provided a mean capacity of 335,840 bits, while the block-based DCT scheme provided only 5,216 bits—a ratio of roughly 64:1. This follows directly from the design: LSB stores one bit in every pixel, whereas the DCT scheme stores one bit in every 8×8 block, i.e., one bit per 64 pixels. Expressed as bits per pixel, LSB reaches up to about 1.0 bpp at full capacity while DCT reaches only about 0.0155 bpp.

Table II lists the maximum capacity of each method for every cover image. Because the capacity depends only on the image dimensions, it is constant across embedding levels. The LSB-to-DCT ratio is close to 64:1 for every image; the small deviation upward (66.0 for the 256×256 *clock* image) arises because the fixed 32-bit length header is proportionally larger relative to the much smaller DCT capacity, an effect that diminishes as the image grows. The table also shows that capacity scales linearly with image size, from about 65 kbit (LSB) for the smallest image to over one megabit for the largest.

TABLE II. MAXIMUM PAYLOAD CAPACITY PER COVER IMAGE

Cover Image	Size (MP)	LSB capacity (bits)	DCT capacity (bits)	Ratio
5112_clock	0.066	65,504	992	66.0
4203_mandrill	0.262	262,112	4,064	64.5
4205_airplane_f16	0.262	262,112	4,064	64.5
4206_sailboat	0.262	262,112	4,064	64.5
4207_peppers	0.262	262,112	4,064	64.5
5210_stream_bridge	0.262	262,112	4,064	64.5
boat512_boat	0.262	262,112	4,064	64.5
5301_male	1.049	1,048,544	16,352	64.1
Mean	0.336	335,840	5,216	64.6

C. *Imperceptibility–Capacity Trade-off*

The central finding of this study concerns the relationship between the two methods when capacity and quality are considered together.

First, at their respective operating points, LSB achieves higher PSNR than DCT. For example, at full capacity LSB yields 51.13 dB versus 44.67 dB for DCT even though LSB is simultaneously carrying roughly 64 times more data. This may appear counter-intuitive, but it has a clear explanation. An LSB change alters a single pixel by at most one gray level. A DCT change, by contrast, modifies a frequency coefficient by up to the quantization step $\Delta = 16$, when the inverse transform is applied, that single coefficient change is spread across all 64 pixels of the block. The net distortion energy introduced per embedded bit is therefore substantially higher for DCT, which is reflected in its higher MSE and lower PSNR despite a far smaller payload.

Second, the two methods do not overlap in bits-per-pixel range, so a direct reading of *Fig. 1* at a common bpp is not possible. To compare them on equal terms, consider equal absolute payloads. The full DCT payload for a 512×512 image is only about 4,064 bits, an LSB fill ratio of roughly 1.5%. Embedding that same payload with LSB and measuring the result directly yields a mean PSNR of 69.19 dB (with SSIM ≈ 1.0000) for the 512×512 images, more than 24 dB higher than the 44.67 dB that DCT produces for the identical payload at its full capacity. Thus, when the comparison is normalized to equal payload, LSB is not merely higher-capacity but also dramatically more imperceptible than the block-based DCT scheme implemented here.

These results should be read in the context of the methods' intended use. The value of transform-domain embedding lies not in raw capacity or spatial-domain fidelity but in its greater resistance to detection by spatial-domain statistical attacks, such as the chi-square and histogram tests. Within the scope of pure imperceptibility and capacity, however, LSB is the stronger performer.

D. *Summary of Observations*

The experiments support three conclusions.

(1) Both LSB and DCT are perceptually imperceptible at all tested embedding levels (PSNR above 44 dB, SSIM above 0.97).

(2) LSB offers approximately 64 times the capacity of the block-based DCT scheme.

(3) For equal payloads, LSB also yields higher PSNR and SSIM, because DCT distributes more distortion per embedded bit across each transformed block. The low variance across eight images of differing sizes indicates that these conclusions are stable.

E. *Practical Implementations*

These measurements can be used for choosing an embedding method. When the goal is to hide a large payload in a single image and the stego image will not be re-compressed or otherwise processed, LSB is the clear choice. It offers both the highest capacity and for any given payload size, the highest fidelity. When the payload is small and the priority is that the hidden data survive routine image handling or resist casual statistical inspection, the lower-capacity DCT scheme is preferable despite its higher per-bit distortion. The results also caution against operating any method near its full capacity, although even a fully embedded image remains visually imperceptible, the companion paper shows that high embedding rates are precisely the regime in which statistical detection becomes easy, so a prudent user keeps the embedding rate well below capacity.

F. *Threats to Validity*

Several factors bound the generality of these results. First, the experiment uses eight standard test images. While their low metric variance is reassuring, a larger and more diverse corpus would strengthen the statistical claims. Second, the analysis is performed on the grayscale channel so the absolute capacities reported for LSB would roughly triple for three-channel color embedding, though the relative comparison with DCT would be unchanged. Third, the DCT scheme studied here embeds a single bit per block in one fixed coefficient. Alternative DCT designs that use multiple coefficients per block would raise its capacity and shift the trade-off. Finally, imperceptibility is assessed with objective metrics (PSNR, SSIM) that approximate, but do not perfectly reproduce human perception. A formal subjective study was outside the scope of this work. None of these factors affects the central, clearly separated finding that LSB dominates DCT on both capacity and equal-payload fidelity.

V. CONCLUSION

This paper presented a controlled experimental analysis of the imperceptibility and payload capacity of LSB and DCT image steganography. Both techniques were implemented from first principles and evaluated on eight standard grayscale test images across five embedding levels using MSE, PSNR, SSIM, and bits-per-pixel metrics. The study found that both

methods keep the stego image visually imperceptible (PSNR above 44 dB and SSIM above 0.97 in every case), that LSB provides roughly sixty-four times the capacity of the block-based DCT scheme, and that LSB additionally achieves higher image quality for an equal payload because each DCT bit perturbs an entire 8×8 block. These results quantify the capacity–imperceptibility trade-off between spatial and transform-domain embedding and provide a reproducible baseline for the companion study, which examines how detectable each technique is under chi-square and histogram steganalysis. Future work includes extending the DCT scheme to embed in multiple coefficients per block to raise its capacity, evaluating color-channel embedding, and testing robustness to lossy compression.

Source Code

https://github.com/echaa0018/Steganography_Analysis.git

Acknowledgment

The author thanks Prof. Dr. Ir. Rinaldi, M.T. as the lecturer for the II4021 Cryptography course and the assistants for their knowledge and guidance. The author also acknowledges the use of open-source scientific Python libraries (NumPy, SciPy, scikit-image, Pillow, pandas, Matplotlib) in the implementation.

References

[1] N. F. Johnson and S. Jajodia, "Exploring steganography: Seeing the unseen," Computer, vol. 31, no. 2, pp. 26–34, Feb. 1998. [Online]. Available: https://doi.org/10.1109/MC.1998.4655281

[2] T. Morkel, J. H. P. Eloff, and M. S. Olivier, "An overview of image steganography," in Proc. 5th Annu. Information Security South Africa Conf. (ISSA), 2005, pp. 1–11. [Online]. Available: http://www.martinolivier.com/open/stegoverview.pdf

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

[4] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, "Digital image steganography: Survey and analysis of current methods," Signal Processing, vol. 90, no. 3, pp. 727–752, 2010. [Online]. Available: https://doi.org/10.1016/j.sigpro.2009.08.010

[5] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th ed. New York, NY, USA: Pearson, 2018. [Online]. Available: https://www.pearson.com/en-us/subject-catalog/p/digital-image-processing/P200000003224

[6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004. [Online]. Available: https://doi.org/10.1109/TIP.2003.819861

[7] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," IBM Systems Journal, vol. 35, no. 3–4, pp. 313–336, 1996. [Online]. Available: https://doi.org/10.1147/sj.353.0313

[8] J. Fridrich, Steganography in Digital Media: Principles, Algorithms, and Applications. Cambridge, U.K.: Cambridge Univ. Press, 2009. [Online]. Available: https://doi.org/10.1017/CBO9781139192903

STATEMENT

I hereby declare that this paper is my own work, not a paraphrase or translation of another person's paper, and is not plagiarism.

Bandung, 19 June 2026

Mahesa Satria Prayata 18223082