Improving the Security of Least Significant Bit (LSB) Steganography in Digital Images Using AES-128 Encryption and Evaluating Its Resistance to Steganalysis Attacks

Aulia Azka Azzahra - 18223131
Information System & Technology Study Program
School of Electrical Engineering and Informatics
Institut Teknologi Bandung, Ganesha Street No. 10 Bandung
E-mail: auliaazka@gmail.com , 18223131@std.stei.itb.ac.id

***Abstract*—This research presents an experimental investigation into enhancing the security of LSB steganography in digital images using Advanced Encryption Standard (AES-128) encryption. The study evaluates image visual quality using PSNR and MSE metrics, data integrity through BER, and system resilience against statistical steganalysis attacks, namely Chi-Square Attack and RS Analysis, under varying payload sizes. The results show that AES-128 pre-processing does not significantly degrade visual quality compared to standard LSB, maintaining PSNR values above 51 dB at a payload of 85,000 characters with completely lossless extraction. In terms of security, the pseudorandom characteristics of AES ciphertext produce a more randomized bit distribution, enabling the system to remain undetected by Chi-Square-based histogram analysis up to an 85,000-character payload. In contrast, standard LSB with plaintext embedding becomes detectable at a payload of 1,000 characters. Additionally, the experiment reveals a statistical dilution effect in global RS Analysis caused by the sequential embedding mechanism. These findings provide practical insights for developing secure digital communication systems that balance data confidentiality, visual quality, and embedding capacity.**

Keywords—LSB Steganography, AES-128, Steganalysis, Chi-Square Attack, Visual Quality, Data Security

I. Introduction

Digital steganography plays a crucial role in modern secret communication by hiding the existence of sensitive data within a cover image, such as a digital image. The Least Significant Bit (LSB) method is the most widely adopted spatial domain steganography technique due to its computational simplicity and its ability to preserve the image's visual quality (imperceptibility). By embedding message bits into the least significant bits of image pixels, the resulting modifications are virtually undetectable by the human visual system. Nevertheless, traditional LSB steganography that embeds plaintext possesses significant security vulnerabilities. The structured, low-entropy bit patterns of text are highly susceptible to forced extraction and detection by modern steganalysis algorithms.

To enhance resilience against detection and forced extraction, securing steganography requires a cryptographic layer. The Advanced Encryption Standard (AES-128) in CBC mode offers a robust solution by transforming the original message into ciphertext prior to the embedding process. Theoretically, the high-entropy (pseudorandom) characteristics of the ciphertext can disguise the data bits as natural noise within the image. However, optimizing the implementation of this combined method still requires a comprehensive empirical understanding of its performance characteristics in practical scenarios, particularly regarding its tolerance level against statistical steganalysis attacks as the embedded payload capacity increases.

This paper specifically addresses this need through a quantitative experimental evaluation of the efficiency and security of AES-128 encryption-based LSB steganography. The specific objectives of this research are to: (1) objectively measure the impact of embedding on the image's visual quality and data integrity using Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), and Bit Error Rate (BER) metrics; (2) empirically test the resilience and detection threshold of this method against histogram anomalies using the Chi-Square Attack across various payload sizes; (3) evaluate the resilience against pixel structural correlation testing through RS Analysis; and (4) discuss the practical trade-offs between payload capacity, visual distortion, and security implications in steganographic systems.

II. Backgrounds and Related Works

A. Cryptography and Advanced Encryption Standard (AES)

Cryptography provides the fundamental foundation for securing data confidentiality. In the context of digital data security, symmetric cryptographic algorithms such as the Advanced Encryption Standard (AES) are frequently utilized due to their high computational efficiency. AES operates on fixed-size data blocks of 128 bits and supports various key sizes, one of which is 128 bits (AES-128).

To avoid vulnerabilities related to pattern analysis prevalent in the basic Electronic Codebook (ECB) mode, modern cryptographic implementations generally employ the Cipher Block Chaining (CBC) mode. CBC mode encrypts text using an XOR operation between the current plaintext block and the previous ciphertext block. This algorithm requires a pseudorandom Initialization Vector (IV) for the first block, ensuring that identical plaintexts will produce entirely different ciphertexts [1]. The high level of entropy in the AES output makes it highly suitable for integration with steganography, as the ciphertext will appear as random noise within the image.

B. Least Significant Bit (LSB) Steganography and Evaluation Metrics

Steganography aims to conceal the existence of secret data. The Least Significant Bit (LSB) spatial method works by replacing the last bit (the 8th bit) of the cover image's pixel intensity values with bits from the secret message [2]. Because the 8th bit has the smallest visual contribution to color representation, manipulation of this bit (an intensity value change of +1 or -1) is generally imperceptible to the human visual system.

To quantitatively measure the impact of embedding on the image's visual quality, two main metrics are used: Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR). MSE measures the average squared error between the original image (C) and the stego-image (S) with pixel dimensions M x N:

$$MSE = \frac{1}{M\,x\,N}\sum_{i=1}^{M}\sum_{j=1}^{N}(C(i,j) - S(i,j))^2 \qquad (1)$$

Meanwhile, PSNR measures the ratio between the maximum possible pixel intensity value (MAX_I, typically 255 for an 8-bit image) and the noise (MSE) on a logarithmic decibel (dB) scale:

$$PSNR = 10\,log_{10}\left(\frac{MAX_I^2}{MSE}\right) \qquad (2)$$

A higher PSNR value indicates better stego-image quality. PSNR values above 40 dB are generally considered to have excellent visual quality and are difficult to distinguish from the original image by human observation [3].

C. Statistical Steganalysis: Chi-Square and RS Analysis

Although LSB preserves the image's visual quality, this method fundamentally alters the statistical distribution of pixel values. Steganalysis attacks are designed to detect these anomalies.

1) *Chi-Square Attack:* This attack exploits the Pairs of Values (PoV) phenomenon. In standard LSB, embedding a message tends to equalize the occurrence frequencies of adjacent even (2k) and odd (2k + 1) pixel values. The X^2 test statistic is used to compare the theoretical occurrence frequency (E_i) with the actual frequency (O_i) on the image histogram:

$$X^2 = \sum_{i=1}^{k}\frac{(O_i - E_i)^2}{E_i} \qquad (3)$$

A significant increase in the X^2 value (approaching a probability/p-value of 1.0) serves as a strong indicator of sequential LSB embedding [4].

2) *RS Analysis:* his is a more sophisticated method that analyzes the spatial correlation between pixels. The RS algorithm divides the image into pixel blocks and classifies them into Regular (R) and Singular (S) groups based on a discriminant function after applying a flipping mask (positive and negative) [2]. In natural images, the percentage of R and S groups on positive masking (R,S) and negative masking (R_m,S_m) are approximately equal. If the difference (diff_R = $|R-R_m|$) exceeds the tolerance threshold, the image is suspected of containing a payload.

D. Related Work

Various aspects of LSB steganography vulnerabilities have been widely documented in previous literature, including its high susceptibility to statistical detection when embedding low-entropy plaintext [4]. Several studies have reported that the integration of cryptography and steganography can enhance system security through the combination of message content protection and message existence concealment [5]. However, most studies have focused more on developing complex pixel modification algorithms rather than empirically analyzing a simple yet practical AES-LSB hybrid scheme.

Furthermore, empirical documentation examining the effect of increased payload capacity on the detection rate by statistical steganalysis methods, particularly in Python-based implementations, remains relatively limited and under-explored. This report expands upon existing knowledge by conducting a direct comparative test between pure LSB and LSB combined with AES-128. This research provides concrete empirical data to evaluate the effectiveness of cryptographic randomization in reducing the detection rate by statistical steganalysis methods such as the Chi-Square Attack and RS Analysis, while also presenting a trade-off analysis between cryptographic security and steganalysis detection resilience in real-image testing.

III. METHODOLOGY

A. System Implementation Details

The implementation of this steganography system, based on the integration of LSB and AES-128, was built using the Python programming language with a modular architecture. The PyCryptodome library was used to manage cryptographic operations, while Numpy and Pillow (PIL) were used for spatial digital image matrix manipulation. The system is divided into three main modules: the encryption/decryption module, the embedding/extraction module, and the steganalysis evaluation module.

In the pre-embedding stage, the message is secured using the Advanced Encryption Standard (AES) algorithm with a 128-bit key length in Cipher Block Chaining (CBC) mode. The secret key is generated from the user's password via the SHA-256 hash function, while the Initialization Vector (IV) is randomly generated using a Cryptographically Secure Pseudo-Random Number Generator (CSPRNG). During the embedding stage, the Least Significant Bit (LSB) algorithm modifies the last bit of pixel intensities sequentially (top-to-bottom). To ensure reliable extraction, the ciphertext is combined with a 64-bit header containing the payload length and a magic number (0xDEADBEEF) as a file integrity validator.

Additionally, the system is equipped with a data processing pipeline subsystem using the pandas library and metric visualization using matplotlib. All test artifacts, including data records (CSV files) and high-resolution graphical outputs (300 DPI), are isolated in a specific directory to maintain the consistency of the research structure.

B. Experimental Design

The empirical testing was conducted through a Command Line Interface (CLI) program executing two main scenarios: Baseline Benchmarking and Payload Variance Testing. The test image used was a natural PNG image with a resolution of 512×512 pixels, providing a theoretical maximum embedding capacity of 98,296 bytes.

1) *Baseline Benchmarking:* This experiment evaluates the direct impact of AES-128 integration on visual quality and data integrity compared to traditional LSB steganography. The test utilized a single 403-character plaintext sample embedded using both methods. The recorded evaluation metrics include Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Bit Error Rate (BER). Additionally, initial Chi-Square Attack and RS Analysis tests were conducted to verify the detection status at a low embedding capacity.

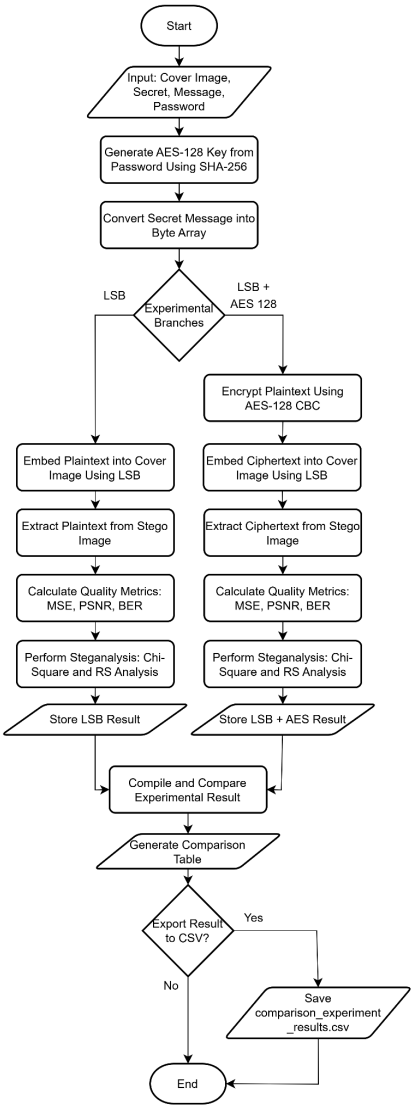

Fig. 1. Performance Comparison Testing Flowchart

2) *Payload Variance Testing :* The subsequent experiment focused on the system's resilience against statistical steganalysis attacks as the payload size increased. Payload size variations were dynamically generated, containing random alphanumeric character combinations, ranging from 100, 1,000, 10,000, 40,000, to 85,000 characters (approaching 86% of the image's maximum limit). At each variation, the system calculated the visual quality metrics and executed steganalysis (Chi-Square and RS Analysis). The primary goal of this mapping was to identify the detection threshold where statistical anomalies begin to be detected (the p-value approaching 1.0 in the Chi-Square test or significant fluctuations in the R and S differentials based on RS Analysis). The comparison results of the conventional LSB method and the hybrid AES-LSB method were then visualized to analyze the trade-off between security and embedding capacity.

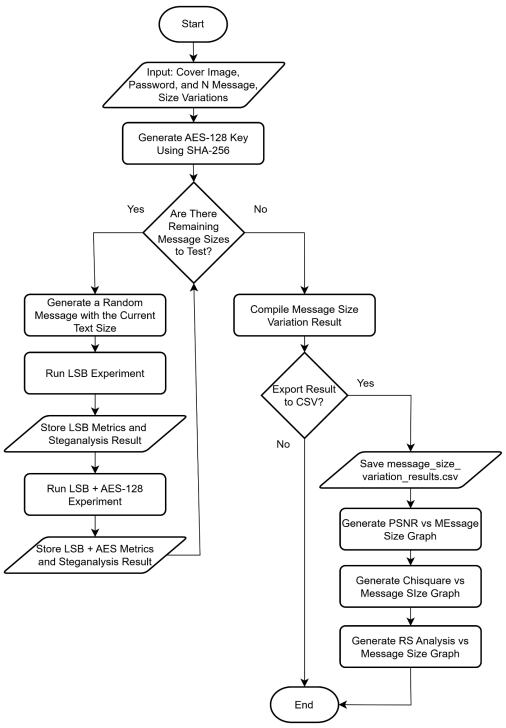

Fig. 2. Payload Scalability Testing Flowchart

IV. Results and Analysis

The testing of this AES-LSB hybrid steganography system was conducted using a PNG cover image with a resolution of 512 × 512 pixels, which has a theoretical maximum embedding capacity of 98,296 bytes. All recorded test data were analyzed in several sections, including baseline benchmarking and payload scalability testing against statistical steganalysis attacks.

A. Baseline Visual Quality and Data Integrity Analysis

The first stage of testing was conducted to determine the fundamental impact of adding the AES-128 encryption layer on the visual quality of the stego-image, as well as to ensure the message integrity remains intact when extracted. This comparative test utilized a fixed secret message sample of 403 characters (~0.4% of the total image capacity). The baseline test results are comprehensively presented in Table 1.

TABLE I. Performance Comparison Results of Standard LSB and LSB + AES-128 (403-Character Payload)

Evaluation Metric	***Standard LSB***	***LSB + AES-128***
Mean Squared Error (MSE)	0.002089	0.002244
Peak Signal-to-Noise Ratio (PSNR)	74.9310 dB	74.6200 dB
Bit Error Rate (BER)	0.0000	0.0000
Chi-Square Statistic Value	109.6159	97.8018
Detection Probability (p-value)	0.9317	0.7545
Chi-Square Detection Status	Not Detected	Not Detected
RS Analysis Detection Status (Global)	Not Detected	Not Detected

Based on Table 1, both testing schemes yielded a Bit Error Rate (BER) of 0.0000. This indicates that all secret data bits can be fully extracted without a single bit experiencing data corruption (lossless extraction), including the successful decryption process in the LSB + AES-128 scheme.

Regarding visual quality, the PSNR values for both methods are in the range of 74 dB, far exceeding the human visual perception threshold limit of 40 dB. The PSNR difference between standard LSB (74.9310 dB) and LSB + AES-128 (74.6200 dB) is relatively very marginal. The slight difference in PSNR values is likely caused by the ciphertext's bit distribution being more random than the plaintext, resulting in slightly different LSB modification patterns. The random nature of these ciphertext bits increases the probability of bit flipping occurring in the image pixels' LSBs compared to plain text bits, whose patterns tend to be more structured.

Furthermore, the payload size used in this test is still very small, at approximately 0.4% of the image capacity. Consequently, the resulting changes in histogram distribution and spatial pixel correlation are not significant enough to produce a strong indication of embedding. Based on the detection applied in this study, both methods are still categorized as undetected in the initial steganalysis testing.

B. Scalability and Visual Quality Distortion Analysis

Subsequent testing focused on scalability analysis to observe the effect of increasing payload sizes on the degradation level of the cover image's quality. The test was conducted by embedding payload variations ranging from a small size (100 characters) to a very large size (85,000 characters, covering ~86% of the maximum image capacity). The visual aspect and integrity test data are summarized in Table 2.

TABLE II. Impact of Payload Size Variations on Image Visual Quality

Size (Characters/Bytes)	***MSE Standard LSB***	***MSE LSB + AES-128***	***PSNR Standard LSB (dB)***	***PSNR LSB + AES-128 (dB)***	***BER***
100	0.000584	0.000696	80.4693	79.7075	0.0
1,000	0.005093	0.005155	71.0614	71.0086	0.0
10,000	0.050733	0.050903	61.0779	61.0633	0.0
40,000	0.203124	0.203740	55.0532	55.0400	0.0
85,000	0.431849	0.432186	51.7775	51.7741	0.0

The data in Table 2 demonstrate a negative correlation between payload size and PSNR value. As the payload size increases, the MSE value increases, causing the PSNR value to drop. This decline follows the logarithmic characteristic of the PSNR metric in response to changes in MSE. The pattern of this visual quality degradation is illustrated more clearly in the graph curve below:

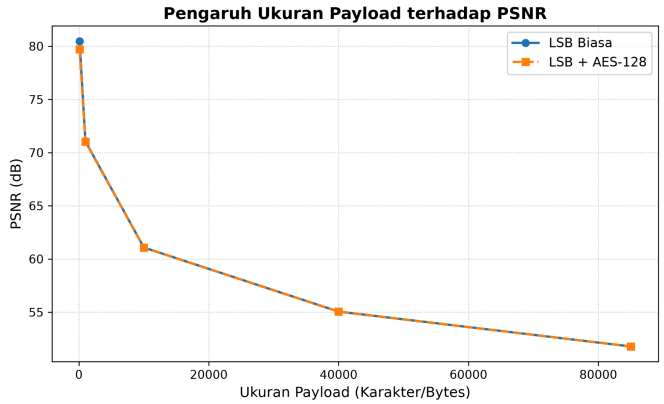

Fig. 3. Graph of Payload Size Impact on PSNR Value

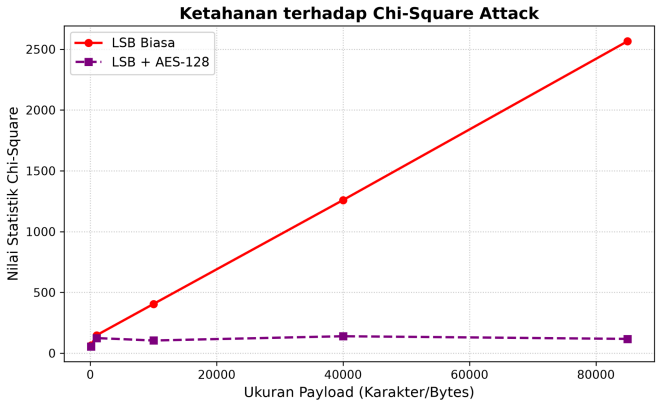

Fig. 4. Graph of System Resilience Comparison against Chi-Square Attack

Based on Figure 3, increasing the payload size consistently causes a decline in PSNR values. This occurs because an increasing number of original pixel LSB bits are modified, elevating the Mean Squared Error (MSE). A crucial phenomenon visible in Figure 3 is that the PSNR degradation curves for the standard LSB method and LSB + AES-128 show nearly identical patterns at every test point. This empirical finding demonstrates that adding the AES-128 encryption security layer imposes no penalty or additional detrimental visual distortion effects on the image. Even at an embedding capacity approaching the image's maximum limit (85,000 characters), the PSNR value remains high at 51.7741 dB, well above the critical threshold of 40 dB. Thus, the imperceptibility aspect of this system is preserved.

C. Resilience Evaluation against Chi-Square Attack

Evaluating resilience against statistical steganalysis attacks serves as the primary parameter to prove the security superiority of the proposed method. The recorded statistical values and detection statuses from the Chi-Square algorithm attacks are presented in Table 3.

TABLE III. CHI-SQUARE ATTACK DETECTION RESULTS ACROSS VARIOUS PAYLOAD SIZES

Size (Characters/ Bytes)	***Chi-Square Stat (Standard)***	***Detection Status (Standard)***	***Chi-Square Stat (AES-128)***	***Detection Status (AES-128)***
100	64.5316	Not Detected	55.3058	Not Detected
1,000	147.9497	**Detected** (Message Present)	124.3970	Not Detected
10,000	404.9538	**Detected** (Message Present)	104.4883	Not Detected
40,000	1259.7948	**Detected** (Message Present)	139.5954	Not Detected
85,000	2565.9817	**Detected** (Message Present)	117.6131	Not Detected

The distinct statistical characteristics between the two methods are visualized through the graph below:

Based on the curve in Figure 4, standard LSB steganography (solid red line) exhibits significant vulnerability. When the payload is increased to just 1,000 characters, the Chi-Square statistic value sharply spikes and produces a DETECTED status with a probability (p-value) reaching 0.9975. This exponential spike in the statistical value continues to climb, reaching 2565.9817 at maximum capacity. This is suspected to occur because plain text characters have a more structured bit distribution compared to ciphertext, resulting in more observable changes in the Pairs of Values distribution on the image's histogram.

Conversely, the LSB + AES-128 method (dashed purple line) proves to have a high level of resilience. Throughout the test, from low loads to maximum capacity (85,000 characters), the statistical curve remains stable, situated within a low-value range, and consistently maintains an UNDETECTED status. This is attributed to the role of AES-128 encryption, which transforms the text structure into ciphertext with a high level of entropy and a more randomized bit distribution than plaintext. When the ciphertext bits are embedded into the image's LSB, the Chi-Square test does not find a statistical anomaly strong enough to indicate the presence of a payload, making the secret data harder to detect via histogram analysis.

D. Detection Anomaly and Statistical Dilution Effect in RS Analysis

The final steganalysis testing method utilized RS Analysis to measure the spatial correlation shift between adjacent pixels via a discriminant function. The visualization of the specific test results on the Red Channel component is displayed in the following graph:

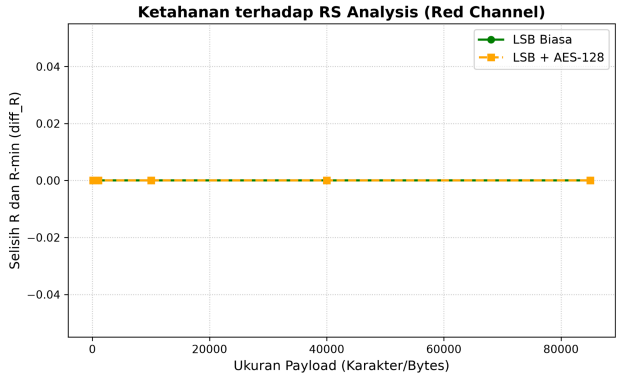

Fig. 5. Graph of Differential Results in RS Analysis (Red Channel)

Based on the plot results in Figure 5, a unique anomaly was found where the differential curve (diff_R) remained at a constant value of 0.0000 for both the standard LSB method and the AES-LSB hybrid method. This constant zero phenomenon indicates that the utilized RS Analysis approach was unable to capture the statistical changes induced by the embedding process in this test configuration. One possible cause is the statistical dilution effect due to the sequential nature of the embedding.

The sequential embedding in this application module operates sequentially (top-to-bottom) using the flatten() function. With a large image resolution size, even a 20,000-byte payload actually modifies only about 20% of the image matrix area at the top. Meanwhile, the remaining 80% of the image area below has not undergone modification. Because the applied RS Analysis algorithm conducts testing globally across the entire image dimension, the dominance of statistical data from the unmodified original pixel area (unaltered pixels) causes the minor anomalies from the local stego area at the top to become less visible.

These experimental findings indicate that the applied RS Analysis implementation is less sensitive in detecting statistical changes resulting from the sequential embedding method in this testing scenario. Further investigation using a Windowed RS Analysis approach or varying other testing parameters is necessary to verify this hypothesis.

V. Conclusion

This research has demonstrated several key characteristics of integrating Least Significant Bit (LSB) steganography with AES-128 encryption, including its preserved visual quality level (imperceptibility), successful data extraction, and resilience against statistical steganalysis attacks. Our findings confirm that applying AES-128 as pre-processing does not introduce additional visual quality degradation to the cover image, where the Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) values show very minimal differences compared to the standard LSB method. The use of AES-128 adds a cryptographic security layer to the steganography system. Furthermore, a Bit Error Rate (BER) of 0.0 indicates that the extraction and decryption processes across all testing scenarios can be executed without data loss (lossless).

Testing payload variations and conducting steganalysis revealed a trade-off between embedding capacity, visual quality, and statistical security. It is essential to implement high-entropy ciphertext encryption to ensure a reliable level of confidentiality in communication environments vulnerable to interception. The safe boundary for payload size must be determined based on the cover image's capacity and texture distribution. Standard LSB with plaintext insertion proved to increase vulnerability to statistical detection, whereas AES-128 encryption potentially strengthens the security aspect by disguising the secret data to resemble natural noise. Therefore, it is recommended to consider the following aspects before configuring and implementing a reliable steganography system:

1) Implementation of Pre-Embedding Encryption (AES-128): Prioritize the use of encryption before embedding messages to balance a large payload capacity with security (reducing the probability of detection). Based on experimental results, the use of AES-128 enhances resilience against Chi-Square Attack detection compared to standard LSB. Given the structured nature of plaintext, standard LSB was detected at a 1,000-character payload. Conversely, AES-128 managed to maintain an undetected status against histogram anomalies even at a maximum payload size of 85,000 characters.

2) Payload Capacity Boundary Management: Although the proposed method demonstrates high resilience against statistical detection, regulating the payload capacity remains necessary to preserve the stego-image's visual quality. Based on the experimental results on the 512 × 512-pixel test image used in this study, it is recommended that the payload size not exceed approximately 85% of the image capacity. Within this range, the PSNR value remains above 50 dB, indicating that the visual quality of the image remains excellent and is difficult to distinguish from the original image.

3) Beware of the Dilution Effect in Steganalysis Testing: Experimental results indicate that local sequential embedding (top-to-bottom) on high-resolution images can trigger a statistical dilution effect when analyzed using global RS Analysis. For communication applications requiring a high level of security, steganalysts must be aware of the weaknesses of this averaging test and adapt the detection method, for instance, by using a Windowed Steganalysis approach (specific testing per matrix block) to avoid false negative results.

These findings provide practical information for security system developers and cryptography practitioners to configure a steganography architecture that offers a balance between strong cryptographic defense, preserved visual quality, and the ability to operate across various image capacities. Potential options for future research include investigating the application of scattered/randomized embedding methods using a Pseudo-Random Number Generator (CSPRNG), comparing resilience against other modern steganalysis algorithms (such as Sample Pair Analysis or Machine Learning-based Steganalysis), and conducting studies to assess the resilience of this LSB+AES architecture against lossy compression schemes like JPEG.

Source code repository at github

https://github.com/auliaazkaazzahra/stego-lsb

Presentation Slides

https://canva.link/w89fcsr0we5zhv4

Video Link at Youtube

https://youtu.be/h0v323ryWyE?si=kt-nv4rI4qPU_LDe

Acknowledgment

The author expresses sincere gratitude to Dr. Ir. Rinaldi Munir, M.T. for his invaluable guidance and inspiring lectures, which motivated the author to delve deeper into the fields of cryptography and steganography. Appreciation is also extended to classmates in the Information Systems and Technology program and laboratory assistants for their constructive discussions and support during the experimental phase of this research.

References

[1] M. Dworkin, "Recommendation for Block Cipher Modes of Operation: Methods and Techniques," National Institute of Standards and Technology (NIST) Special Publication 800-38A, 2001. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-38a.pdf

[2] J. Fridrich, M. Goljan, dan R. Du, "Detecting LSB steganography in color and gray-scale images," IEEE MultiMedia, vol. 8, no. 4, hal. 22-28, Okt.-Des. 2001. https://staff.emu.edu.tr/alexanderchefranov/Documents/CMSE492/Fridrich%20Steganalysis%202001.pdf

[3] O. O. dkk., "Image Steganography Method using LSB and AES Encryption Algorithm," CEUR Workshop Proceedings, Vol. 4016, 2024. https://ceur-ws.org/Vol-4016/paper3.pdf

[4] A. Westfeld dan A. Pfitzmann, "Attacks on Steganographic Systems," dalam Information Hiding (IHW 1999), Lecture Notes in Computer Science, vol. 1768, Springer, 1999. https://users.ece.cmu.edu/~adrian/487-s06/westfeld-pfitzmann-ihw99.pdf

[5] Purwanto, A. Marjuni, E. Z. Astuti, C. A. Sari, N. Rijati, P. N. Andono, and M. K. Sarker, "Enhancing image cryptography using Fibonacci AES-LSB," TELKOMNIKA (Telecommunication Computing Electronics and Control), 2024. https://telkomnika.uad.ac.id/index.php/TELKOMNIKA/article/download/26078/11990

STATEMENT

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 19 Juni 2026

Aulia Azka Azzahra (18223131)