iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79830

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Steganography Detection System for Cyber Forensics

Rhutu Abhijeet Khire[1], Prachi Krishnat Jadhav[2], Harshada Santosh Barge[3], Prof. Mujawar S. A[4]
Department of Computer Engineering, Karmaveer Bhaurao Patil Polytechnic, Satara

Abstract: ***The swift expansion of digital media has led to steganography becoming a significant challenge in cyber forensics, allowing malicious individuals to hide sensitive information within seemingly harmless carrier files. Steganography is defined as the practice of embedding confidential data within another data stream, enabling it to be sent to its destination without arousing suspicion. Among the various carrier types, images are most commonly used due to their high data capacity and widespread availability. Several techniques have been developed for embedding information within digital images, with the Least Significant Bit (LSB) method being one of the most widely adopted. This paper presents the technique for detecting steganography, allowing forensic investigators to uncover and scrutinize digital assets. The proposed method utilizes statistical analysis approaches, including Chi-square and RS steganalysis, to automatically detect concealed payloads. Experimental results demonstrate that the approach achieves high accuracy across diverse range of datasets, effectively distinguishing between "clean" and "stego" images, even at low embedding rates are used.***

Keywords: LSB, Cyber forensic, chi-square, RS steganalysis

I. INTRODUCTION

Effective communication is a crucial aspect of modern life, as the swift evolution of communication technologies has made the sharing of digital media-such as images, audio, and video-vital. This progress also requires careful attention to data security, copyright issues, and the threats posed by cyber criminals who exploit these public platforms for hidden communication. Steganography involves concealing information in a manner that makes hidden signals imperceptible. It is a method for embedding messages within a suitable carrier, like images, audio, or video file. This allows the message to be transmitted to the recipient without anyone being aware that it contains a secret communication. Johnson et al. proposed the magic triangle model to assess steganographic techniques. This model highlights three key components: embedding data rate, inaudibility, and robustness. A higher data rate signifies that a greater amount of information can be encoded within the carrier. Robustness pertains to resilience against steganographic attacks. There exists a trade-off among these three factors, complicating efforts to improve one without negatively impacting the others. Figure 1 illustrates the relationship among three components. The substitution of the least significant bit (LSB) is a commonly used steganographic technique. The least important bits are utilized for both embedding and extraction, leading to minimal distortion of the image.

Fig.1 Johnson's Magic Triangle Model

The term Steganography comes from the Greek terms “stegos” and “grafia,” which indicates “cover” and “writing” respectively. Thus, it might be defined as “covered writing”. Image steganography is generally categorized into four main types: spatial domain, frequency domain (also known as transformation domain), spread spectrum, and model-based steganography. In spatial domain techniques, data bits are inserted directly into the pixel values of the images. Model-based steganography relies entirely on the statistical model of the cover image. A widely used method for embedding data bits is the Least Significant Bit (LSB) substitution. The fundamental concepts include cover image, stego image, imperceptibility, capacity, and security. The cover image servers as the carrier image in steganography for concealing text data, while the payload is embedded data. Together, the cover image and payload form the stego image. Capacity indicates the volume of data that can concealed within a cover image, measured in bits per pixel (bpp). The stego image should not contain any recognizable artifacts resulting from message embedding. Such artifacts could be exploited by a third party to reveal the existence of a concealed message. If a third party can consistently identify which images hold hidden messages, the effectiveness of the steganographic technology is compromised. Although not the most secure approach of embedding data in photos, LSB steganography tools are already widely used. As a result, detecting images with hidden messages produced by LSB embedding is critical in terms of effectiveness, accuracy, and reliability. In LSB replacement steganography, the LSB of every pixel is substituted with a secret data bit, enabling the concealment of confidential information across the image. The hidden capacity (HC) will be one bit per pixel. If we wish to hide more bits, we can make the swap up to two or three LSBs. In rare circumstances, we can expand this notion up to four LSBs to hide a very high number of bits. However, if we go up to 4 LSBs, distortion will be substantial, making it easy to detect using various detection algorithms. The LSB replacement is easy and identifiable using regular-singular (RS) analysis. RS analysis is a steganalysis mechanism that correctly detects the LSB substitution. In steganography, the unique image is referred to as the cover image, while the message-implanted image is known as the stego image. Steganography consists of three components: the carrier, message, and key. Steganos translated to “covered” or “secret,” while graphy refers to writing or drawing.

Least significant bit (LSB) steganography is a widely used and simple technique for hiding information within a cover image. In the LSB method, the bits of the secret message substitute the bits of the cover image. To enhance capacity, data can be embedded in two or more LSBs; however, this may compromise the quality of the stego image. Basic LSB steganography can be identified through statistical techniques like RS and Chi-square analysis. LSB steganography can be identified through statistical techniques like RS and Chi-square analysis. For embedding, a pair of pixels’ servers as a unit, where the first pixel’s LSB carries one bit of information, and the other bit is conveyed by a function of the both pixels. Consequently, this revised method allows for the payload to remain unchanged, even though cover image is altered less often.

Steganography seeks to conceal signals within a medium, rendering them undetectable, in contrast to cryptography, which permits outsiders to identify, intercept, and alter messages. Unlike cryptography, steganography not only encrypts messages but also obscures the very existence of the communication. Steganography serves various purposes, such as managing copyright, enhancing engine robustness, and incorporating personal data into images. It also includes video-audio synchronization, secure transmission of confidential information, television broadcasting, and more. In this study, we present improved steganalysis algorithms that utilized the most dependable thinly-spread LSB steganography detectors currently accessible, emphasizing the application of grayscale Bitmaps as cover images. A successful steganographic system requires a cover medium that blends in with the stego-object. This study intends to: a) improve steganalysis statistics for LSB steganography, b) employ large image libraries to provide experimental proof, and c) identify the bit rate upper limits that prevents LSB steganography detection.

II. IMAGE BASED STEGANOGRAPHY

Embedding a message within an image necessitates two files. The cover image is a seemingly harmless image that is utilized to conceal information. The second file holds the message that contains information intended to be hidden. Memorandums can take the form of plain text, cipher text, images, or embedded within bits streams. When distributed, the cover image along with cover message forms a stego images. A stego-key, which acts as a type of password, can be employed to conceal and subsequently decode a message. Steganography software often recommends lossless 24-bit images. The most common of these are BMP files. When a hidden message is inserted into a cover image using a certain method, the output is known as a stego image. The stego image seems visually identical (or nearly identical) to the source image, reducing suspicion during transmission.

To improve security, a stego-key (akin to password) is commonly utilized during the embedding and extraction processes. This key determines how and where the data is concealed within the image, guaranteeing that only authorized users can access the hidden information. Without the proper stego-key, retrieving the concealed message is extremely difficult. The Least Significant Bit (LSB) approach is a popular technique for image steganography. In this method, the least significant bits of pixel values are altered to store

the secret data. Because these bits have little effect on overall image quality, the alterations are virtually invisible.

A pixel value of 11001010 can be altered to 11001011 to incorporate one bit of secret information. Minor modifications have little effect on the image, hence LSB is a preferred option. Understanding how data is embedded in images is critical to discovering it. Chi-Square Analysis and RS Steganalysis are techniques created specifically for identifying statistical and structural changes introduced during embedding. Chi-Square Analysis reveals anomalies in pixel value distributions due by LSB changes. RS Steganalysis investigates pixel smoothness and noise patterns to better uncover buried data.

III. CHI-SQUARE AND RS STEGANALYSIS

Chi-Square steganalysis is a statistical technique used to uncover concealed information within digital images by analyzing the distribution of pixel values. In a natural image, pixel values follow predictable distribution. Conversely, when data is embedded through methods such as Least Significant Bit (LSB) replacement, the distribution becomes altered. The Chi-Square test compares the observed frequency of pixel values to the expected frequency in an unchanged image. If the discrepancy between these frequencies is considerable, it suggests the presence of concealed data. This approach is especially useful for identifying simple LSB embedding, in which pixel values are updated in a deterministic manner. The primary advantage of Chi-Square steganalysis is its simplicity and low computer complexity, making it ideal for fast analysis of huge datasets. It is less successful against more advanced embedding techniques, such as LSB matching or adaptive steganography, because the modifications are less predictable and more equally dispersed.

RS (Regular-Singular) steganalysis is more advanced technique that addresses the limitations of simple statistical methods. It operates by examining the structural features of pixel clusters within an image rather than their frequency distribution. This method divides the image into small groups of pixels and uses a discriminating function to determine the smoothness or noise level inside each group. This analysis divides pixel groups into two categories: Regular (R) and Singular (S). Regular groups have smooth patterns, whereas singular groups have more noise or inconsistency. During the embedding process, particularly in LSB-based approaches, the natural equilibrium between regular and singular groups is upset. RS steganalysis can discover the presence of hidden data and estimate the embedding rate by making controlled alterations, such as flipping the least significant bits and analyzing the ensuing changes in group categorization. Unlike Chi-Square analysis, RS steganalysis may detect more subtle changes induced by steganographic techniques, making it more reliable and accurate. It works well with both grayscale and color images and provide a deeper understanding of the image's structure. However, this method is more computationally demanding and necessitates careful parameter selection, such as group size and discrimination functions. Furthermore, while RS steganalysis is successful against LSB substitution, it may still have difficulties in detecting very sophisticated or adaptable embedding systems.

Chi-Square and RS steganalysis are both critical components of modern cyber forensic systems. Chi-Square analysis is a quick and effective techniques to uncover evident statistical irregularities, but RS steganalysis allows for a more extensive and reliable evaluation of pixel structures. When these strategies are coupled, they improve overall detection capabilities, allowing for the more accurate and confident identification of concealed information in digital images.

IV. RELATED WORK

The most widely used image formats include GIF, JPEG, and PNG and many existing steganographic techniques are designed for these format. However, some approaches utilize the Bitmap format (BMP) due to its simple and uncompressed data structure. The LSB substitution technique generally does not result in an increase in file size. However, depending on the volume of information being concealed, the file may undergo considerable distortion. Steganographic tools such as StegHide, S-tool, and Stegnos use LSB substitution for data hiding.

Least significant bit (LSB) steganography is a straightforward method for embedding information in cover images. Increasing the number of bits for embedding improves the data capacity but may degrade image quality, creating a trade-off between payload size and visual imperceptibility. Basic LSB method can be detected using statistical approaches such as RS and Chi-square analysis. Some improved techniques modify embedding strategies to reduce detectable changes while maintaining the same payloads. To perform the embedding, a pair of pixels are utilized as a unit. The least significant bit (LSB) of the first pixel carries one bit of information, while the second pixel's LSB carries the other bit. The revised technique permits embedding of the same payload with minimal alterations to the cover image. Mohammad's approach involves storing a variable number of bits in each channel (R, G, or B) of a pixel depending on its color value. The lower color component used to store more bits, thereby enhancing the hiding capacity of the cover image. The indicator is randomly selected from the three channels based on the pre-agreed key between the sender and recipient. This method maintains the histogram form for all three channels.

G. Karthigai improved the LSB insertion method to protect against statistical attacks like RS and Chi-square. This was achieved by first encrypting and hiding data along The technique preserves the statistical and visual characteristics of the cover image by embedding the secret message in the sharper edge areas, determined by a threshold based on the gradients and the size of the message. Huang introduced a method for LSB replacement method that combines segments based on the size of the secret message and the difference between neighboring pixels in the cover images.

Steganography and steganalysis have garnered significant attention within the field of information security, owing to their use clandestine communication and cyber forensics. Throughout the years, numerous researchers have introduced both statistical method and structural methods for detecting hidden information within digital images. Among the most essential and widely utilized techniques for identifying Least Significant Bit (LSB) based steganography are Chi-Square analysis and RS (Regular-Singular) steganalysis. Initial investigations in steganalysis focused on statistical methods aimed at examining the distribution of pixel values. The Chi-Square attack, which utilizes Pearson's Chi-Square statistical test, was one of the early effective approaches for reveling hidden data in images. It computes the disparity between the observed and anticipated frequency distributions for pixel pairs. Research indicates that this approach is highly effective in identifying LSB replacement schemes, where the embedding process equalizes the frequencies of pixel pair. Nonetheless, a major drawback noted in the literature is that Chi-square analysis is limited to frequency distribution, overlooking the spatial relationships between pixels, which reduces its efficiency when compared to more advanced embedding methods.

V. METHODOLOGY

The proposed Steganography Detection System for Cyber Forensic employs combination of statistical and structural steganalysis techniques to accurately detect hidden information buried in digital images. The system uses a multi-stage pipeline that includes data collecting, preprocessing, feature extraction, detection, classification, and forensic analysis. The combination of Chi-Square and RS steganalysis improves detection accuracy and robustness against a variety of LSB-based embedding methods.

A. Data Acquisition

The initial involves gathering input photos for analysis. The system accepts:

- Digital photos in formats including BMP, PNG, and JPEG
- Both clean (cover images) and possible stego images

Images can be sourced form:

- Local storage
- Network transfers
- Digital Forensic Datasets

These images are the major data source for finding concealed information.

B. Data Preprocessing

Before analysis, the input images are preprocessed to maintain consistency and increase detection performance. The preprocessing steps include the following:

- Image Normalization (Resizing and Format Standardization).
- Color space conversion (RGB to grayscale if necessary).
- Noise reduction through filtering techniques.
- Pixel value extraction for statistical computing.

This phase ensures that all images are in consistent format, ready for further processing.

The Steganography Detection System uses a series of procedures to efficiently analyze photos and locate hidden data.

Algorithm steps:

1. Start
2. Upload/Input images into the system.
3. Validate image format and size.
4. Preprocess the image (resize, normalize if needed).
5. Use detecting techniques:

- Perform the Least Significant Bit (LSB) analysis.
- Apply the Chi-square statistical test.
- Perform an RS steganalysis.

6. Extract characteristics and investigate pixel behavior.
7. Calculate the statistical deviations and anomalies.
8. Calculate a risk score based on analysis.

9. Classify the image as safe or suspicious.
10. Display output results and analysis details.
11. Ask the user if they want to analyze another image.
12. If so, repeat the process.
13. If not, quit the system.
14. End

Flowchart (Text Representation)

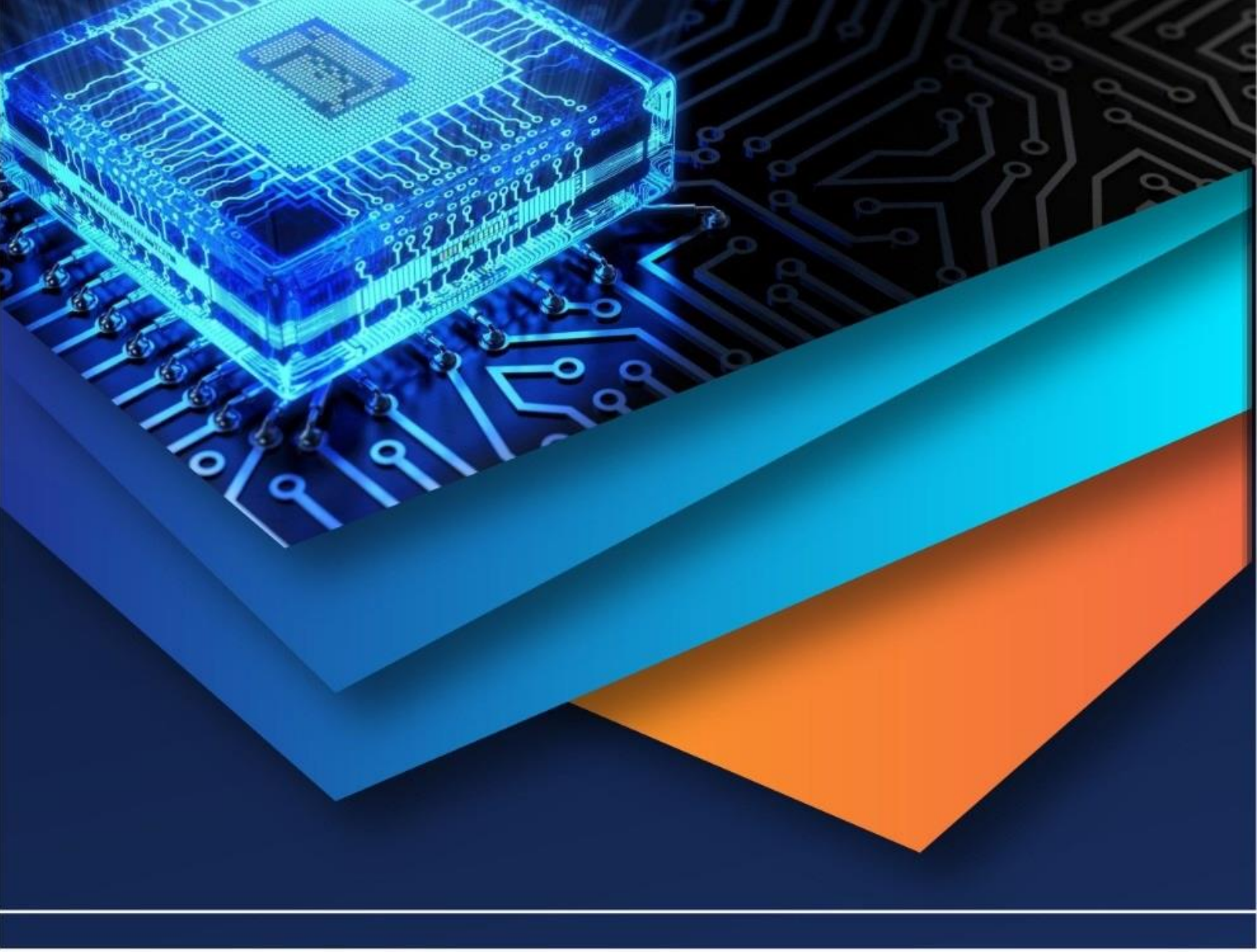
Start
↓
Upload Image
↓
Validate Input
↓
Preprocess Image
↓
Apply Detection Techniques
↓
Analyze Pixel & Statistics
↓
Calculate Risk Score
↓
Classify Image
↓
Display Result
↓
Continue? → Yes → Repeat
→ No → End

VI. CONCLUSION

The Steganography Detection System was effectively developed and implemented to identify concealed information within digital images, addressing the pressing need for automated analysis using statistical and analytical methods in the realm of cyber forensics amid the rising prevalence of covert communication and data hiding techniques. This approach employs multiple detection methods, including Least Significant Bit (LSB) analysis, Chi-square testing, and RS steganalysis, to evaluate pixel-level alterations and identify a typical patterns resulting and data embedding, thereby enhancing detection accuracy and reducing false positives, while the incorporation of a risk score mechanism further aids in clearly indicating the probability of concealed data and simplifying result interpretation. The initiative emphasizes usability by providing a straightforward interface for users to input images and receive prompt analysis results, thereby reducing the necessity for manual inspections, conserving time, and enhancing efficiency in cyber forensic investigations; furthermore, the system is meant to be scalable and may be enhanced with modern technologies such as machine learning to increase its performance in the future. The steganography detection system addresses the shortcomings of current methodologies by providing a dependable, efficient, and automated means of uncovering concealed information, thereby enhancing cybersecurity, aiding in digital evidence analysis, and mitigating the potential misuse of steganography for harmful intents.

REFERENCES

[1] Wu, Da-Chun, and Wen-Hsiang Tsai. "A steganographic method for images by pixel-value differencing." Pattern Recognition Letters 24, no. 9 (2003): 1613-1626.
[2] Johnson, Neil F., and Sushil Jajodia. "Exploring steganography: Seeing the unseen." Computer 31, no. 2 (1998): 26-34.
[3] Fridrich, J., Goljan, M., Du, R.: Steganalysis based on JPEG compatibility. In Tescher, A.G., Vasudev, B., Bove, Jr, V.M., eds.: Multimedia Systems and Applications IV. Volume 4518 of Proc. SPIE. (2002) 275–280
[4] C.-K. Chan and M. Cheng, “Hiding data in images by simple LSB substitution,” Pattern Recognition, vol. 37, no. 3, pp. 469-474, 2004.
[5] T. Morkel, J.H.P. Eloff, and M.S. Oliver, “An overview of image steganography,” in Proc. ISSA, pp. 1-11, 2.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY