Chi-Square and Histogram-Based Steganalysis Detection of LSB and DCT Steganographic 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*—Steganalysis is the counterpart of steganography, the task of deciding whether an image carries hidden data. This paper evaluates two classical statistical steganalysis techniques—the chi-square attack on pairs of values and a histogram pair-of-values analysis against images produced by the Least Significant Bit (LSB) and Discrete Cosine Transform (DCT) embedding methods studied in the companion paper. Both detectors were implemented in Python and applied to a set of eighty stego images and eight covers derived from eight standard grayscale test images, with high-entropy payloads embedded at five rates (10%, 25%, 50%, 75%, and 100% of capacity). Detection performance was measured with the threshold-free Area Under the ROC Curve (AUC) per embedding rate, together with an operating-point analysis using a data-driven decision threshold. The results show that detectability rises monotonically with the embedding rate for LSB at full capacity the chi-square attack reaches an AUC of 1.000 and the histogram analysis 0.984, while at a 10% rate both fall to about 0.56. Both detectors are markedly weaker against DCT (AUC at most 0.70 at any rate), confirming that spatial-domain statistical attacks largely fail to detect frequency-domain embedding. The chi-square attack consistently outperforms the histogram analysis.**

Keywords—steganalysis; chi-square attack; histogram analysis; LSB; DCT; ROC; AUC

I. Introduction

Steganography conceals the existence of a message by hiding it inside an innocuous cover such as a digital image. Its adversary is steganalysis, the discipline of detecting whether a given object contains hidden information. Where the steganographer seeks imperceptibility, the steganalyst seeks any statistical irregularity that betrays embedding. The two fields advance together, each new embedding method invites a new attack, and each successful attack motivates a more careful embedding scheme.

Reliable steganalysis matters wherever covert communication is a concern. Digital-forensics examiners use it to recover evidence of hidden data exchange, enterprise and network-security tools use it to flag images that may carry exfiltrated data or concealed malware command-and-control instructions, and law-enforcement and content-moderation workflows use it to detect illicit material hidden inside innocuous-looking files. In every case the analyst must decide from the image alone and without the embedding key, whether a payload is present with the same blind-detection setting evaluated in this paper. Quantifying how detection reliability depends on the embedding method and rate is therefore of direct practical value because it tells a defender which steganographic threats current statistical tools can and cannot catch.

The companion paper analyzed two embedding techniques, LSB substitution in the spatial domain and DCT coefficient embedding in the frequency domain, and showed that both are visually imperceptible, at every tested payload the stego images were indistinguishable from their covers under PSNR and SSIM. Visual imperceptibility, however, does not imply statistical undetectability. LSB substitution in particular leaves a well-known statistical fingerprint. It tends to equalize the frequencies of pairs of pixel values that differ only in their least significant bit. This regularity is the basis of the chi-square attack introduced by Westfeld and Pfitzmann, one of the earliest and most influential steganalysis methods and of related histogram-based analyses.

This paper asks a direct question, given the stego images produced by the LSB and DCT methods, how reliably can simple statistical steganalysis detect them, and how does that reliability depend on the amount of data embedded? To answer it, two detectors, a chi-square attack and a histogram pair-of-values analysis are implemented and evaluated across a sweep of embedding rates on a common image set.

The contribution of this work is an experimental characterization of detection performance as a function of embedding rate and embedding method. Rather than reporting a single accuracy figure, the study measures the Area Under the ROC Curve (AUC) at each embedding rate, which is independent of any decision threshold and therefore gives a fair picture of how separable cover and stego images are. The analysis yields three findings: detectability grows monotonically with payload and becomes near-perfect for LSB at full capacity, low-rate embedding is difficult to detect, and both spatial-domain detectors are largely blind to DCT embedding. The very property that makes LSB attractive (high capacity) is also what makes it most detectable when heavily used.

The remainder of the paper is organized as follows. Section II reviews steganalysis and the two detection techniques. Section III describes their implementation, the dataset, and the evaluation methodology. Section IV presents and analyzes the results. Section V concludes.

II. Theoretical Background

A. Steganalysis

Steganalysis is fundamentally a binary classification problem. An image is either a cover (no hidden data) or a stego object (carrying a hidden message/file). A detector computes a score reflecting the likelihood of embedding and compares it to a threshold to reach a decision. Because the steganalyst rarely knows the embedding rate or key in advance, blind or statistical steganalysis relies on general regularities of natural images that embedding disturbs. The detectors studied here are passive and target the spatial-domain statistics of pixel values.

Steganalysis methods are commonly classified along two axes. By goal, an attack may be passive, merely deciding whether a payload exists ((as in this study) or active, additionally estimating the payload length or attempting to recover or disrupt it. By scope, a targeted attack is designed against a specific embedding algorithm and exploits its particular statistical fingerprint, whereas a blind or universal attack uses generic image features or machine learning to flag anomalies regardless of the embedding method. The two detectors evaluated here are passive and targeted, both are tailored to the pair-of-values signature that LSB substitution leaves behind. This specialization is also why, as the results show, they are far less effective against the structurally different DCT scheme, whose distortion does not produce that signature.

B. The Chi-Square Attack

The chi-square attack of Westfeld and Pfitzmann exploits the structure of LSB substitution. Consider the image histogram and group intensity values into pairs of values (PoV) that differ only in the least significant bit: *(0, 1), (2, 3), …, (2i, 2i + 1)*. In a natural cover image the two members of a pair generally have unequal frequencies, $h(2i) \neq h(2i + 1)$. LSB substitution with a balanced (random) message redistributes pixels within each pair, driving the two frequencies toward their common mean. The attack therefore tests, with a chi-square statistic, how close the observed frequencies are to the value expected under embedding. A high resulting probability indicates that the histogram pairs have been equalized, i.e., that embedding is likely. Applying the test to successive regions of the image yields a per-region embedding probability, averaging these gives an overall estimate that also reflects the fraction of the image that has been embedded.

Fomally, let h□ denote the frequency of intensity value j in the image histogram, and group the 256 values into k pairs *(2i, 2i + 1)*. Under the hypothesis that the least-significant-bit plane has been overwritten with a balanced random message, the two members of each prair tend toward equality, so the expected frequency of each member is their average:

$$y_i = (h_{(2i)} + h_{(2i+1)}) / 2 \quad (1)$$

The attack measures the discrepancy between the observed frequency of the even member, $h_{(2i)}$, and this expected value using Pearson's chi-square statistic, summed over the k pairs whose expected frequency is non-zero:

$$\chi^2 = \Sigma_i (h_{(2i)} - y_i)^2 / y_i \quad (2)$$

Under the embedding hypothesis this statistic follows a chi-square distribution with $v = k - 1$ degrees of freedom. The probability that the image carries a payload is obtained from the complement of the chi-square cumulative distribution function $F_{(\chi^2,v)}$:

$$p = 1 - F_{(}\chi^2, \nu_{)}(\chi^2) = 1 - \gamma(\nu/2, \chi^2/2) / \Gamma(\nu/2) \quad (3)$$

Where γ is the lower incomplete gamma function and Γ the gamma function. A stego image whose pairs have been equalized produces observed frequencies close to y_i, hence a small χ^2 and a probability p near 1, whereas a natural cover with strongly unequal pairs produces a large χ^2 and a probability near 0. This p is exactly the per-region embedding probability that the detector computes and averages over the eight image regions.

C. Histogram Pair-of-Values Analysis

The histogram analysis used here is a simpler, related detector built on the same pair-of-values idea. For each pair it measures the relative difference between the two frequencies, $|h(2i) - h(2i + 1)| / (h(2i) + h(2i + 1))$, and averages this over all pairs. A cover with strongly unequal pairs yields a large mean difference, whereas a fully embedded image, whose pairs have been equalized, yields a small one. The detector reports a score that increases as the pairs become more equal, so that higher scores indicate more probable embedding. Because it uses only the global histogram and no statistical model of the expected counts, it is weaker than the chi-square test but serves as an informative baseline.

D. Evaluation Metrics

A binary detector is characterized by its true-positive rate (TPR, the fraction of stego images correctly flagged) and false-positive rate (FPR, the fraction of covers wrongly flagged). Both depend on the chosen decision threshold. The Receiver Operating Characteristic (ROC) curve plots TPR against FPR as the threshold varies, and the Area Under this Curve (AUC) summarizes detection quality independently of any single threshold. An AUC of 1.0 denotes perfect separation of cover and stego, while 0.5 denotes a detector no better than chance. Because AUC requires no threshold, it is the primary metric in this study, threshold-dependent accuracy, TPR, and FPR are reported as a secondary operating-point analysis.

Two properties make AUC especially suitable here. First, it is invariant to the decision threshold, so it measures how well a detector ranks stego images above covers rather than how well one particular cutoff performs an important distinction when the best threshold is unknown in advance. Second, it is

insensitive to class imbalance, which matters because the evaluation set contains ten times as many stego images as covers. A threshold-based accuracy can look deceptively high simply because the stego class dominates the count. Reporting AUC separately for each embedding rate, as in Section IV, further prevents the easy high-rate cases and the hard low-rate cases from being averaged into a single, uninformative figure.

III. Design and Implementation

A. System Overview

The detectors and the evaluation pipeline were implemented in Python 3 using NumPy and SciPy, reusing the same embedding library developed for the companion paper so that the stego images analyzed here are produced by exactly the methods described there. The steganalysis routines form a side-effect-free module, a separate experiment script generates the stego set, scores every image, and computes the detection metrics. All randomness is seeded for reproducibility, and the detectors operate on the grayscale representation of each image.

B. Chi-Square Detector

The chi-square detector computes the pair-of-values histogram and evaluates the chi-square statistic comparing each observed even-valued frequency to the mean of its pair. The image is divided into eight successive regions, the test is applied to each, producing a per-region embedding probability, and the detector reports the mean of these probabilities as its embedding-probability score. The regional formulation follows the original sequential attack and makes the score robust to partial embedding, since regions that contain the embedded payload register high probabilities even when the rest of the image does not.

C. Histogram Detector

The histogram detector computes the global pair-of-values histogram and returns the equalization score described in Section II-C. One minus the mean relative frequency difference over all pairs. The score lies in [0, 1] and increases as the histogram pairs become more equal.

D. Dataset and Stego Generation

Eight standard grayscale test images served as covers, stored in lossless TIFF format and ranging from 256×256 to 1024×1024 pixels. For each cover, stego images were generated with both the LSB and the DCT method at five embedding rates: 10%, 25%, 50%, 75%, and 100% of each method's capacity yielding ten stego images per cover. Including the eight covers, the evaluation set comprised 88 images (8 covers and 80 stego objects).

A methodological point deserves emphasis. the embedded payloads were high-entropy uniform random bytes rather than natural-language text. This reflects realistic practice, in which secret data is encrypted or compressed before embedding and therefore appears random, and it is also a precondition for the chi-square attack, whose equalization signature depends on the embedded bits being balanced. Low-entropy payloads would partially mask this signature and understate the detectors' true capability.

E. Evaluation Methodology

Every image was scored by both detectors, and the scores were analyzed in two complementary ways. First, the AUC was computed separately for each embedding rate and method, comparing the eight cover scores against the eight stego scores at that rate, this threshold-free measure is the primary result. Second, an operating point was established for each detector by selecting the single decision threshold that maximizes Youden's J statistic (TPR − FPR) over the pooled data, and the resulting accuracy, TPR, and FPR were recorded. Because this threshold is selected on the same data it is evaluated on, the operating-point figures are best read as an optimistic, illustrative summary rather than an estimate of performance on unseen images. The AUC remains the fairer characterization.

IV. Testing and Result Analysis

A. Detection AUC by Embedding Rate

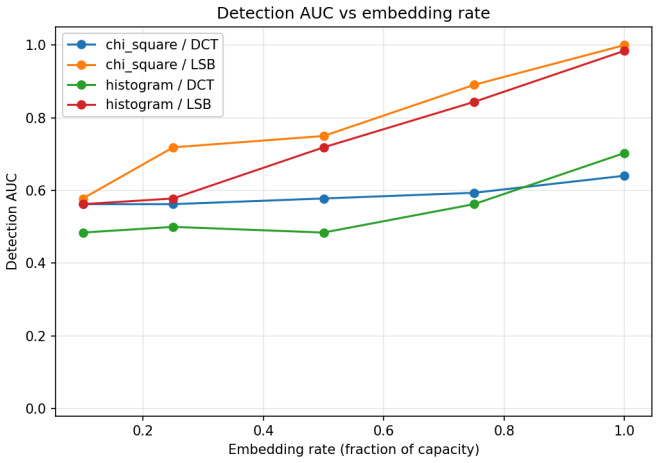
Table I reports the AUC of each detector against each method at every embedding rate. This is the central result of the study.

TABLE I. Detection AUC by Detector, Method, and Embedding Rate

Embedding Rate	Chi-square LSB	Chi-square DCT	Histogram LSB	Histogram DCT
10%	0.578	0.563	0.563	0.484
25%	0.719	0.563	0.578	0.500
50%	0.750	0.578	0.719	0.484
75%	0.891	0.594	0.844	0.563
100%	1.000	0.641	0.984	0.703

Three patterns are evident. First, detectability increases monotonically with the embedding rate for every detector–method combination. The more data is hidden, the more the pixel statistics are disturbed and the easier detection becomes. Second, LSB at high embedding rates is highly detectable, at full capacity the chi-square attack achieves perfect separation (AUC 1.000) and the histogram analysis very nearly so (0.984). Third, low-rate embedding is hard at 10% both detectors fall to roughly 0.56, only marginally better than chance, because a lightly embedded image perturbs the histogram too little to be distinguished from natural variation. Fig. 1 shows these AUC curves.

Fig. 1. Detection AUC versus embedding rate for each detector and method.



B. *LSB versus DCT Detectability*

The two methods differ sharply in how easily they are detected. Against LSB the detectors are strong at high rates, but against DCT they remain weak at every rate, the chi-square AUC for DCT never exceeds 0.641 and the histogram AUC never exceeds 0.703. This is the expected result. The chi-square and histogram attacks both target the spatial-domain pair-of-values signature that LSB substitution produces directly. DCT embedding modifies frequency coefficients; the inverse transform spreads each change across an entire 8×8 block, perturbing pixel values only indirectly and without the characteristic pairing of adjacent intensities. Consequently the spatial detectors register only a weak signal, which grows slightly at the highest DCT rate but never approaches reliable detection.

This finding links the two papers directly. The companion paper showed that LSB offers far greater capacity than DCT. The present paper shows that this capacity is a double-edged property, when LSB is used at high embedding rates it is also the most detectable, whereas DCT, despite its low capacity and higher per-bit distortion, is effectively invisible to these classical spatial attacks.

C. *Operating-Point Analysis*

Table II reports the threshold-dependent performance at the Youden-optimal operating point. Fig. 2 and Fig. 3 show the per-rate true-positive rate at this threshold and the mean detector score, respectively.

TABLE II. OPERATING-POINT PERFORMANCE (YOUDEN'S J THRESHOLD)

Detector	Variables				
	Threshold	**Accuracy**	**TPR**	**FPR**	**Pooled AUC**
Chi-square	0.0026	0.841	0.875	0.500	0.703
Histogram	0.9370	0.466	0.425	0.125	0.642

These binary figures are deliberately less flattering than the per-rate AUC, and they expose the practical difficulty of fixing a single threshold. The chi-square detector attains 84.1% accuracy and an 87.5% true-positive rate, but only by accepting a 50% false-positive rate. The histogram detector, whose optimal threshold sits at the high score of 0.937, holds its false-positive rate to 12.5% but then detects only 42.5% of stego images. The contrast with Table I is explained by two effects: a single threshold cannot simultaneously suit the easy (high-rate) and hard (low-rate) cases pooled together, and several covers produce anomalously high scores, as discussed next.

Fig. 2. True-positive rate versus embedding rate at the chosen threshold.

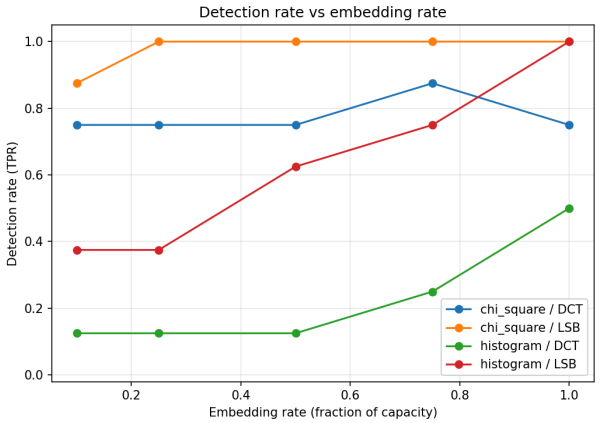
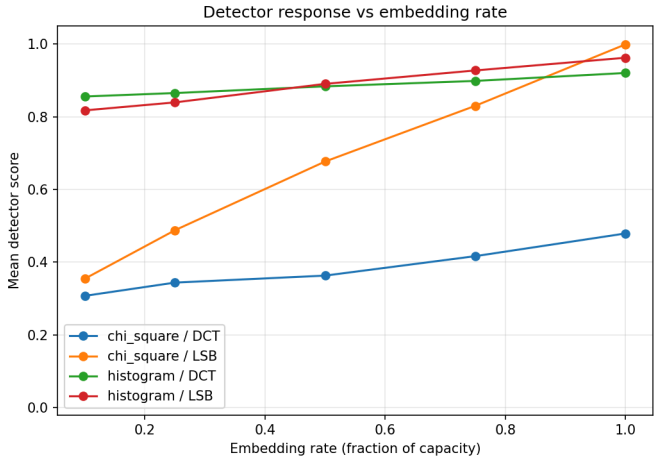


Fig. 3. Mean detector score versus embedding rate.



D. *False Positives on Cover Images*

Table III lists the detector scores for the eight unembedded covers, which ideally should all be low. They are not

uniformly low, and the pattern is instructive. For the chi-square detector, five covers score near zero (correctly), but the highly textured *mandrill* image scores 0.985 and *peppers* scores 0.689, strong false positives. Textured, high-frequency images have a noisy least-significant-bit plane whose pair-of-values frequencies are already nearly equal, so the attack mistakes natural noise for embedding; this is a documented limitation of the chi-square method. The histogram detector fares worse: seven of the eight covers score above 0.86 because natural image histograms are smooth and their adjacent intensity bins are already nearly equal, which the equalization score interprets as embedding. Only the *stream_bridge* image, whose histogram has strongly unequal pairs, scores near zero for both detectors.

TABLE III. DETECTOR SCORES FOR COVER (UNEMBEDDED) IMAGES

Cover Image	Chi-square Score	Histogram Score
4203_mandrill	0.985	0.931
4205_airplane_f16	0.001	0.936
4206_sailboat	0.512	0.945
4207_peppers	0.689	0.937
5112_clock	0.188	0.923
5210_stream	0.000	0.000
5301_male	0.000	0.902
boat512_boat	0.000	0.863

These false positives are the reason the operating-point accuracy in Table II is modest even though the high-rate AUC in Table I is excellent. They illustrate a real limitation of simple statistical steganalysis, the very image characteristics that the detectors exploit (near-equal pair frequencies) also occur naturally in textured or smooth covers, producing confusions that no single threshold can eliminate.

E. Summary of Observations

The experiments support four conclusions:

(1) Detection performance rises monotonically with the embedding rate.

(2) LSB embedding at high rates is highly detectable near-perfect AUC at full capacity while low-rate embedding is difficult.

(3) Both spatial-domain detectors are largely blind to DCT embedding (AUC at most 0.70 at any rate).

(4) The chi-square attack consistently outperforms the histogram analysis, and both suffer false positives on certain natural covers, which limits their accuracy at any fixed decision threshold.

F. Practical Implications

These results carry a clear message. The threshold-free AUC, not a fixed-threshold accuracy, is the meaningful basis for comparing steganalysis detectors, and the operating point should be chosen according to the relative cost of the two error types in the deployment. Where missing a hidden payload is the greater risk, a low threshold maximizes detection at the price of more false alarms. Where a false accusation is costly, a higher threshold is safer but lets lightly embedded images slip through. The findings also delimit what these classical tools can do, they reliably catch only heavily embedded LSB images and are effectively defeated by DCT embedding and by low embedding rates.

G. Threats to Validity

Several limitations bound these conclusions. The cover set comprises eight images, so the true- and false-positive rates are coarsely quantized and the decision threshold is selected on the same data it is measured on which makes the operating-point figures optimistic rather than predictive of unseen images. The payloads are high-entropy, representing the encrypted-data scenario most favorable to the chi-square attack. Lower-entropy payloads would weaken its equalization signature and lower detection. Only two classical targeted detectors are studied. Stronger modern methods like RS analysis, sample-pair analysis, and learning-based steganalysis would likely detect lower embedding rates and resist the cover false positives observed here. These limitations qualify the absolute numbers but not the main point which is that detection improves with payload and spatial attacks do not transfer to the frequency domain.

V. CONCLUSION

This paper evaluated two classical statistical steganalysis techniques, chi-square pair-of-values attack and a histogram pair-of-values analysis against LSB and DCT stego images using detection AUC across a sweep of embedding rates as the primary threshold-free metric. Detectability was found to increase monotonically with the embedding rate, reaching near-perfect separation for LSB at full capacity (chi-square AUC 1.000, histogram 0.984) and falling to little better than chance at a 10% rate. Both detectors proved largely ineffective against DCT embedding, with AUC never exceeding 0.70, confirming that spatial-domain statistical attacks do not capture the frequency-domain signature of DCT steganography. An operating-point analysis with a data-driven threshold revealed that false positives on textured and smooth covers limit the accuracy attainable with any single threshold, a genuine limitation of these simple detectors. Taken together with the companion paper, the results show that LSB's high capacity is also its weakness under heavy use, whereas DCT trades capacity for statistical undetectability against these attacks. Future work includes evaluating stronger detectors such as RS analysis and sample-pair analysis, enlarging the cover set to stabilize the operating-point estimates, and testing detection against adaptive embedding that deliberately preserves pair-of-values statistics.

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] N. Provos and P. Honeyman, "Hide and seek: An introduction to steganography," IEEE Security & Privacy, vol. 1, no. 3, pp. 32–44, May 2003. [Online]. Available: https://doi.org/10.1109/MSECP.2003.1203220

[3] A. Westfeld and A. Pfitzmann, "Attacks on steganographic systems," in Information Hiding, LNCS 1768. Berlin, Germany: Springer, 2000, pp. 61–76. [Online]. Available: https://doi.org/10.1007/10719724_5

[4] 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

[5] 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

[6] T. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters, vol. 27, no. 8, pp. 861–874, 2006. [Online]. Available: https://doi.org/10.1016/j.patrec.2005.10.010

[7] 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

[8] A. D. Ker, "Steganalysis of LSB matching in grayscale images," IEEE Signal Processing Letters, vol. 12, no. 6, pp. 441–444, Jun. 2005. [Online]. Available: https://doi.org/10.1109/LSP.2005.847889

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, 1 June 2026

Mahesa Satria Prayata 18223082