# Analysis of different Steganography Algorithms

**Deepika Bansal**

Department of Computer Science and Engineering, The NorthCap University, Gurugram, Haryana

**Ashu**

Department of Computer Science and Engineering, Dronacharya College of Engineering,Gurugram, Haryana

## ABSTRACT

*Steganography is the technique to hide the important data in the various file formats like image, audio, video, text, etc. In this paper, the image steganography is considered and the performance of various steganography tools is discussed. The analysis of images obtained from different steganography tools is performed on the basis of classification accuracy and histogram analysis. Our main aim is to analyze the most robust and high payload steganography tool.*

***Keywords***

***Steganography, LSB, DCT, Steganography Tools, Feature Extraction, SVM Classifier, Classification Accuracy, Histogram Analysis***

## INTRODUCTION

Steganography is the technique for stowing away the secret messages in the other medium like images, text, video, audio, etc. [1]. The two files are required for hiding the secret information. The first one is the cover and the second consists of the secret message. The secret message can exist in any one of the forms - plain text, cipher text, or image. On hiding the secret message in cover medium, a stego file is obtained. In this paper, the image steganography is considered, in which the images are used for cover medium.

Cover Image + Secret Message = Stego Image

There are various steganographic techniques used for hiding the secret information. In the old days, wax covered tablets, hidden tattoos, invisible inks, microfilms, microdots, null ciphers were used for steganography purpose. The most commonly used image steganography techniques are (i) (i) Spatial Domain & (ii) Transform domain, shown in Fig. 1.

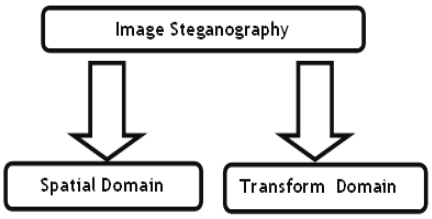



**Fig -1:** Image Steganography Techniques

In the spatial domain technique, secret bits are embedded straightforwardly in the cover image. The Least Significant Bit Insertion (LSB) is most commonly used in Spatial domain. In LSB, the least significant bits are used to hide the secret bits of message. LSB Replacement and LSB Matching are the two types of LSB. The least significant bits of the carrier are replaced by the secret message bit directly in the LSB Replacement technique. But in LSB Matching, if the least significant bit of the cover pixel is increased or decreased by one if it is different from the message bit.

The algorithm for embedding the text message considering LSB technique is described below [4]:

1. The cover image and the secret message to be communicated are taken as input from the user.
2. Then the secret message is transformed into binary.
3. The lsb of each pixel of cover medium is calculated.
4. The lsb of cover image is being replaced by the bits of binary text message.
5. Then the stego image is formed and saved.

The algorithm for retrieving text message:-

1. The stego image is considered as the input.
2. The lsb of each pixel of stego image is calculated.
3. Then retrieve the least significant bits and convert each 8 binary bits into character.

In the transform domain [5], the significant parts of the cover image are used to hide the secret bits. Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) and Discrete Fourier Transform (FFT) are used in transform domain. In DCT technique, the JPEG image format uses a *discrete cosine transformation* to transform 8 x 8 pixel blocks of the image into 64 DCT coefficients each. The DCT coefficients *F(u,v)* of an 8 x 8 block of image pixels *f(x, y)* are given by the following equation:

$$F(u,v) = \frac{1}{4} C(u)C(v) \left[ \sum_{x=0}^{7} \sum_{y=0}^{7} f(x,y) \cos\frac{(2x+1)u\pi}{16} \cos\frac{(2x+1)v\pi}{16} \right]$$

where,

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} & , \text{if } u \leq 0 \\ 1 & , \text{if } u > 0 \end{cases}$$

In [6], the author had compared two steganography tools, PQ and F5 using SVM as a classifier for DCT features showing the reliability of PQ more than F5. In [7], SVM and Neural Network Pattern Recognition tool is used for features extracted from DCT domain on three steganography algorithms – nsF5, PQ and Outguess. It is shown that Outguess is more reliable for steganography. A study on various steganography techniques is carried out in [8]. PSNR value and histogram analysis is done for analyzing cover and stego images. A steganographic technique for hiding secret data is proposed in [9]. The quantized DCT coefficients are used for hiding the secret information. The proposed method provides a high information hiding capacity & increases the security also. In [10], the author had used

GLCM based features for steganalysis for three steganography methods nsF5, JP Hide & Seek and PQ using 3 classification algorithms i.e. J48, SMO and Naïve Bayes.

## STEGANOGRAPHY TOOLS USED

- **F5**

F5 offers a large steganographic capacity and withstands visual and statistical attacks. For improving the embedding efficiency of F5 matrix encoding is implemented. The permutative straddling is employed for spreading out the changes uniformly over the whole stego image. The coefficients are permuted and then embedded in the permuted sequence [11].

- **JP Hide & Seek**

JP Hide & Seek (JPHS) [12]is a Windows program for hiding the secret message inside a jpeg file and to recover the secret message in the jpeg file designed by Allan Latham. The lsb of the discrete cosine transform coefficients is modified in the JPHS tool. In [13], the steps are described by the author for testing the JPHS program. $X^2$-test &Stegdetect are used for detection.

- **nsF5 (no-Shrinkage F5)**

The nsF5 (no-shrinkage F5) is an enhanced version of F5 [14]. It was introduced in 2007. The wet paper codes are used for lessening the negative effect of shrinkage in nsF5 [15].

- **Outguess**

Outguess [16] is a universal steganographic tool. The secret message is embedded in the redundant bits of cover image. The Outguess algorithm [17] relies on the data specific handlers for extracting the redundant bits and writes them back after modifying them. It uses a PRNG to select DCT coefficients [18].

- **Perturbed Quantization**

In PQ [19], an information reducing operation involving quantization such as down sampling, lossy compression or A/D conversion are used for hiding the secret information in the cover image. The main aim of Perturbed Quantization is to achieve minimal distortion and high efficiency [20].

## PROPOSED METHODOLOGY

In this section, the classification framework for distinguishing the various steganography tools is described on the basis of 274 features extracted from DCT domain. SVM is used for feature classification.

The main steps included in the proposed framework shown in Fig. 2 can be described as follows:

(A) Collect the image samples.

(B) Samples are divided into training and test image samples.

(C) DCT domain features are extracted.

(D) Train the SVM classifier.

(E) Classifier is used from (D) for finding the test images.

(F) Find the accuracy of test images.

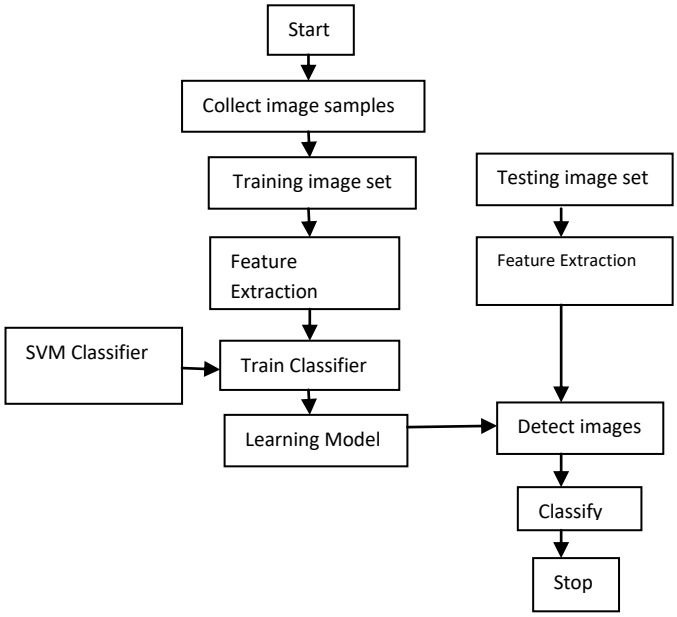



**Fig -2:** The framework for obtaining the classification accuracy of Steganographic algorithms

## FEATURE EXTRACTION IN DCT DOMAIN

All the experiments performed in this study, demonstrates that the steganography calculations like Jsteg, PQ, F5 change the measurable conveyance of DCT coefficients at various recurrence of the picture, in spite of the fact that F5 stays away from the "point-pair" phenomena. It is sure that the relationship inside DCT coefficients will get aggravated when image is installed with message. In light of this hypothesis, measurable model for DCT coefficients were given by Pevny and Fridrich [21]. The different highlights extricated for our work from DCT area are as given underneath in Table 1.

**Table -1: Extended DCT Features Set with 193 features**

| **Feature** | **Dimensionality** |
|---|---|
| Global Histograms | 11 |
| AC Histograms | 5 X 11 |
| Dual Histograms | 11 X 9 |
| Variation | 1 |
| Blockiness | 2 |
| Co - occurence | 5 X 5 |
| Markov Features | 9 X 9 |

## SUPPORT VECTOR MACHINE

We utilizethe support vector machine (SVM) for comparative studyof F5, JPHS, nsF5, Outguess and PQintermsoftheiraccuracy usingthe features extractedin sectionIV. SVMsare used forcalculatingthemaximum margin between linearly separable data. Themarginis calculated by solving an optimization problem. The feature space is modified using kernels inordertoallow fittingof non-linearly separable data. Individualkernels likepolynomial andradial basisfunction are used totransformthefeaturespaceand the performance of kernelis basedoncharacteristics ofthe source data.Gaussianradial-basedkernelswereselected becauseoftheir goodperformance.

## EXPERIMENT RESULTS AND ANALYSIS

All the experiments are implemented in Matlab R2010a. We have downloaded 1000 natural images from www.1000pictures.com [22].We have used five steganography methods F5, nsF5,PQ, JPHS and Outguess, for obtaining five sets of stego images. In the results presented below, for all stego image sets and corresponding cover image sets, 750 images are used to train SVM classifier, and the 250 images are used to test in each set. LibSVM [23] tool using radial basis function was employed for classification. Table II shows the classification results for classification accuracy obtained using SVM. The accuracy of nsF5 is 52.4% and 64.28% is of JPHS. As can further be observed SVM classifies F5 with an accuracy of almost 98.6% and PQ with an accuracy of 90% whereas Outguess images are detected with an accuracy of 56.8%, hence making nsF5 steganography algorithm more robust to steganalytic attack using the features extracted in DCT domain.

The cover image and stego images obtained using different steganography algorithms are shown in Fig. 3 to Fig. 8 and the corresponding histograms are shown in Fig. 9 to Fig. 10.

**Table -2: Classification Accuracy Obtained From Various Steganography Tools**

| **Tools** | **Classification Accuracy** |
|---|---|
| F5 | 98.6 |
| JPHS | 64.28 |
| nsF5 | 52.4 |
| Outguess | 56.8 |
| PQ | 90 |

**Fig -3:** Cover Image

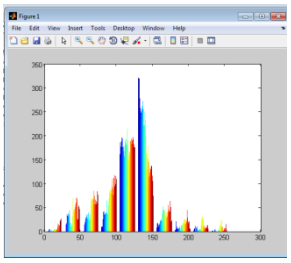
**Fig -4:** F5Stego Image

**Fig -5:** JPHS Stego Image

**Fig -6:** nsF5 Stego Image

**Fig -7:** Outguess Stego Image

**Fig -8:** PQ Stego Image



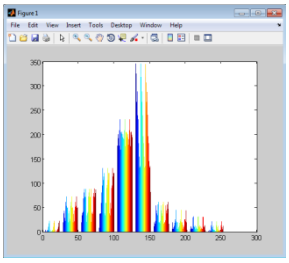
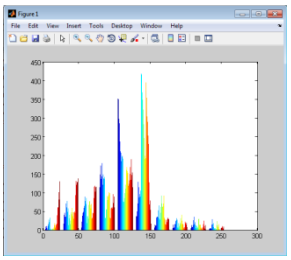


**Fig -9:** Histogram of Cover Image **Fig -10:** Histogram of F5 Stego Image



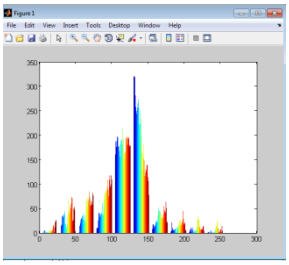
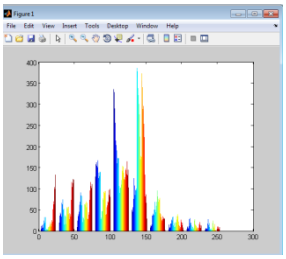


**Fig -11:** Histogram of JPHS Stego Image **Fig -12:** Histogram of nsf5Stego



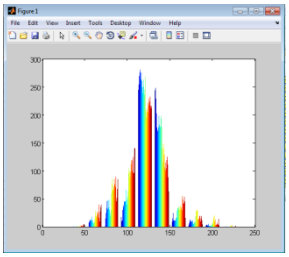


**Fig. 13** Histogram of Outguess Stego Image Fig.14 Histogram of PQ Stego Image

It can be easily seen that the noise content is not perceived by a human eye for the nsF5, from the histogram. Although, there is visible noise in all the histograms obtained from other steganographic algorithms.

## CONCLUSION

Steganography is used for hiding secret information in other file formats, to prevent the existence of data from the attacker. The Image steganography is discussed in this paper using five steganography tools for embedding the secret data in an image. The analysis of stego images using the classification accuracy and histogram analysis is performed. As it can be observed from the above analysis, we can conclude that nsF5 is more robust to steganalytic attacks, as it is yielding the better accuracy and histogram in comparison to other steganography tools discussed in this paper. In future, we shall extend our work to comparative study of other steganography algorithms using more efficient classification techniques. We shall also study the method of feature selection and optimization.

## REFERENCES


[1] N.F.Johnson, S.Jajodia, Exploring Steganography: Seeing the Unseen IEEE Computer 31(2) (1998)26-34.

[2] Kh. Manglem Singh, S.Birendra Sigh and L. ShyamSundar Singh, Hiding Encrypted Message in the Features of Images, IJCSNS, Vol.7 No. 4, April 2007, pp 302-307.

[3] W.-N. Lie and L.-C. Chang, Data hiding in images with adaptive numbers of least significant bits based on human visual system, in Proc.,IEEE Int. Conf. Image Processing, 1999, Page(s): 286–290.

[4] K.B.Shiva Kumar, K.B. Raja, R.K.Chhotaray, SabyasachiPattnaik, "Coherent Steganography using Segmentation and DCT", IEEE-978-1-4244-5967-4/10/$26.00 ©2010.

[5] S. Katzenbeisser and F. A. P. Petitcolas, Information Hiding Techniques for Steganography and Digital Watermarking. Norwood, MA: Artech House, 2000.

[6] Chhikara, R., &Bansal, D. Classification of PQ Stego-Images and F5 Stego-Images using SVM.

[7] Bansal, D., &Chhikara, R. (2014, February). Performance evaluation of steganography tools using SVM and NPR tool. In 2014 Fourth International Conference on Advanced Computing & Communication Technologies (ACCT) (pp. 483-487). IEEE.

[8] DeepikaBansal, Rita Chhikara, A Study on Steganography Techniques, International Journal of Engineering Research & Technology (IJERT)Vol. 3 Issue 2, February – 2014

[9] Bansal, D., &Chhikara, R. (2014). An improved DCT based steganography technique. International Journal of Computer Applications, 102(14).

[10] Chhikara, R. R., &Bansal, D. (2014, September). GLCM based features for steganalysis. In Confluence The Next Generation Information Technology Summit (Confluence), 2014 5th International Conference- (pp. 385-390). IEEE.

[11] A. Westfeld, High capacity despite better steganalysis (F5 – a steganographic algorithm). In I. S. Moskowitz, editor, Information Hiding, 4th International Workshop, volume 2137 of Lecture Notes in Computer Science, pages 289–302, Pittsburgh, PA, April 25–27, 2001. Springer-Verlag, New York.

[12] JPHS: http://linux01.gwdg.de/~alatham/stego.html

[13] JPHS:http://io.acad.athabascau.ca/~grizzlie/Comp607/programs.htm

[14] J. Fridrich, T. Pevný, and J. Kodovský, Statistically undetectable JPEG steganography: Dead ends, challenges, and opportunities. In J. Dittmann and J. Fridrich, editors, Proceedings of the 9th ACM Multimedia & Security Workshop, pages 3–14, Dallas, TX, September 20–21, 2007.

[15] J. Fridrich, M. Goljan, and D. Soukal, Wet paper codes with improved embedding efficiency. IEEE Transactions on Information Forensics and Security, 1(1):102–110, 2006.

[16] Nielsprovos, www.outguess.org.

[17] Provos, N. Defending Against Statistical Steganalysis. Proc. 10th USENIX Security Symposium. Washington, DC, 2001.

[18] N. Provos, P. Honeyman, Hide and seek : an introduction to to steganography, IEEE Security and Privacy 1 (3) (2003)32–44.

[19] J. Fridrich, M. Goljan, and D. Soukal, Perturbed quantizationsteganography. ACM Multimedia System Journal, 11(2):98–107,2005.

[20] Abbas Cheddad, Joan Condell, kevin Curran, paulMcKevitt, Digital image steganography: Survey and analysis of current methods, Signal Processing 90 (2010) pp 727-752.

[21] T. Pevny and J. Fridrich, Merging Markov and DCT features for multi-class JPEG steganalysis. In E.J. Delp and P.W. Wong, editors, Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX, volume 6505, pages 3 1 - 3 14, San Jose, CA, January 29 - February 1, 2007.

[22] Image Source : www.1000pictures.com

[23] LibSVMToolBox Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm/.