Steganalysis Using LSB-Focused Statistical Features

Mudhafar M. Al-Jarrah
Middle East University
Jordan
maljarrah@meu.edu.jo

Zaid H. Al-Taie
Middle East University
Jordan
zaid.altaie@yahoo.com

Abdelrahman Abuarqoub
Middle East University
Jordan
Aabuarqoub@meu.edu.jo

ABSTRACT

In this paper, we present an image steganalysis model with a new texture feature set that is designed to take into consideration the pattern of embedding locations in a cover image. The chosen feature set in based on statistical texture features of images including gray level co-occurrence matrix (GLCM), Entropy, and additional statistical image features that can discriminate between clean and stego images. The guiding principle in choosing the feature set elements is that steganography techniques embed secret data in the right half-byte of an image's bytes, the least significant bits, to avoid perceptible visual distortion that can result from embedding in the left half-bytes. Therefore, the proposed features are applied to 2-LSB, 3-LSB and 4-LSB bit planes of a cover image as well as the full-bytes. For the experimental work, the grayscale single-channel image format was chosen for cover images, and we used the public BossBase1.01 dataset which consists of 10,000 PGM images. The selected classifier was the Support Vector Machine algorithm as implemented in MATLAB. Embedding of data in the cover images was based on 2LSB and 4LSB spatial domain schemes. The feature vectors of clean images, 2LSB stego images and 4LSB stego images, 10,000 each, were analyzed. The detection accuracy results of the validation phase was 99.41% for the combined clean and 4LSB images, and 99.02% for the clean and 2LSB stego images. The paper ends with conclusion and suggestions for applying the proposed model to multi-channel images, and for dealing with alternative steganography schemes.

CCS CONCEPTS

• **Security and privacy → Software and application security; Systems security → File system security**

KEYWORDS

Information hiding, steganalysis, steganography, GLCM, entropy, texture features; grayscale image; dual classification.

ACM Reference format:

M. Al-Jarrah, Z. Al-Taei , A. Aboarqoub. 2017.
Steganalysis Using LSB-Focused Statistical Features.
In *Proceedings of ICFNDS '17, July 19-20, 2017, Cambridge, United Kingdom*, 5 pages. DOI: 10.1145/3102304.3109814

ICFNDS '17, July 19-20, 2017, Cambridge, United Kingdom © 2017 Association for Computing Machinery.

ACM ISBN 978-1-4503-4844-7/17/07...$15.00
http://dx.doi.org/10.1145/3102304.3109814

1 INTRODUCTION

Steganalysis is the art and science of detecting secret messages hidden inside cover media, using steganography [1]. The objective of steganalysis is to gather any evidence about the presence of hidden data [2].

Steganalysis has applications in computer forensics, insider's threats detection, cyber warfare, monitoring criminal and terrorist communications over the internet and gathering evidence for investigations. Research in steganalysis has focused mainly on
detecting hidden data in image carriers of various formats, color and grayscale, lossy compressed, lossless compressed and un-compressed. Steganalysis techniques are grouped into two types: targeted and blind. Targeted steganalysis attempts to detect specific embedding schemes which can give high detection accuracy, but it can fail to detect hidden data if the embedding scheme is unknown [3]. Alternatively, blind steganalysis can be considered a general scheme for detecting different types of steganography. However, as blind steganalysis can detect a wider range of steganography techniques, it is expected to be less accurate in detecting stego images or in miss-detecting a clean image as stego.
More advanced steganalysis techniques attempt to extract additional information about the hidden message than just its presence. Quantitative steganalysis is an estimator of the number of embedding changes introduced by a specific embedding operation [4]. The number of embedding changes can help in estimating the message length, hence quantitative steganalysis can be an important forensic tools.

Feature-based steganalysis aims to provide blind staganalysis, for example the research in [5, 6]. The key factor in blind steganalysis is the feature set, the different statistics that are extracted from the carrier image which can discriminate between clean and stego images.

The pioneering paper by Haralick[7] presented an approach for analyzing the texture of an image through extracting statistics from the co-occurrence matrix. The paper proposed the gray level co-occurrence matrix (GLCM) feature set which consisted of texture descriptors extracted from the co-occurrence matrix, including contrast, energy, homogeneity correlation and entropy. The GLCM feature set was used in experimental work which showed high detection performance [8].

Many methods of image texture analysis have been developed and experimentally evaluated, with the aim of improving the detection accuracy of steganalysis, based on the different metrics and features of image texture characterization. Some of the high performing methods include histogram of oriented gradients (HOG) [9], gradient location and orientation histogram (GLOH) [10], region covariance matrix (RCM) [11], gray level co-occurrence matrix (GLCM) [7], local binary patterns (LBP) [12], and color gradient co-occurrence matrix (CGCM) [13].

The work by Ker [14] experimentally evaluated the GLCM method, using a large dataset of grayscale images. The work evaluated stego images that were generated by different steganography schemes, taking into consideration variations in embedded data size.

In this paper we present a blind steganalysis feature set that includes GLCM and other image texture statistics, with focus on the LSB part of an image's bytes, the right half-byte (RHB). The reason for paying attention to the right half-byte is that it is the area where most embedding occurs, regardless of the embedding scheme.

2 LSB-FOCUSED FEATURE SET

The proposed feature set is designed to focus on texture properties of the right half-byte (RHB) of the 8-bit channel of the grayscale image, thereby giving attention to the part of a cover image which is most often used in embedding, regardless of steganography scheme. The 2LSB, 3LSB and 4LSB bit-planes are treated as sub-images when the features are calculated. The main components of the feature set is the Gray Level Co-Occurrence Matrix (GLCM) ( ), the Entropy ( ) and additional statistical metrics that aims to contribute to the discrimination between stego and clean images.

Table 1 shows the list of proposed feature set elements. The correlation coefficient (CC) between the two halves of a byte (LHB vs. RHB) is an indication of the change in association between the two halves after embedding in the right half-byte.

Table 1: Feature Set Elements

Feature Element	Feature Description
CC-LR	Correlation coefficient between LHB and RHB
CVR	Coefficient of variation of right half-bytes
GLCM-B	Contrast, Correlation, Homogeneity Energy, of full bytes
GLCM-R	Contrast, Correlation, Homogeneity, Energy, of RHB
GLCM-3	Contrast, Correlation, Homogeneity, Energy, of 3LSB
GLCM-2	Contrast, Correlation, Homogeneity, Energy, of 2LSB
Entropy-R	Entropy of RHB
Diff-R	Average of absolute difference between successive RHB

The coefficient of variation (CV), which is the ratio of standard deviation to the average, is calculated for the right half vertical slice of the image, as CVR. It is assumed that the coefficient of variation of the right half-byte will change as a result of embedding. The GLCM features are calculated for the entire image, i.e. the column of full bytes, as GLCM-B. The GLCM-R features are calculated for the right half vertical slice of the image, which consists of a column of right half-bytes. The GLCM-3 and GLCM-2 features are calculated for the 3LSB and 2LSB vertical slices of the image. The GLCM for 1LSB vertical slice was excluded as it did not show any discriminating effect during the experiments. The Entropy function is calculated for the right half vertical slice of the image, as Entropy-R. The average of the absolute difference between successive right half-bytes, Diff-R, focuses on the effect of changes to the right half of the image, as a result of embedding.

3 THE DATASET

The selected cover images dataset in this work is the BOSSBASE1.01 dataset [15], which consists of 10,000 grayscale images, in PGM format. All the images have the same dimensions of 512x512, which gives image size of 256 KB. Fig. 1 shows a sample from the dataset. Two secret images with different sizes are used for embedding.

Figure 1: A sample cover image from the BOSSBASE1.01 dataset [15]

For 4LSB embedding, the image LightHouse.jpg, shown in Fig. 2, is used. The image size is 125 kb which occupies about 50% of the hiding capacity of the cover images using the 4LSB scheme.

Figure 2: The secret image LightHouse.jpg, size 125 kb, converted to JPG from Kodak Lossless True Color Image Suite [16]

A smaller secret image is used for 2LSB embedding, in which 25% of the cover pixel area is utilized for embedding. Fig. 3 shows the image Lena_496.jpg, with image size is 63 kb.

Figure 3: The secret image Lena_496.jpg, size 63 kb, converted to JPG from Gonzales standard image set [17]

4 EXPERIMENTS

After selecting the image dataset, the experimental work involved four phases: stego images generation through steganography embedding, feature extraction, classification through training and validation testing, and results analysis.

4.1 Stego images generation

Stego images were generated from the clean cover images dataset, using the 2LSB and 4LSB schemes. The embedding process implemented a spatial domain LSB replacement in which two or four LSB bits were replaced in the cover image's bytes. The two secret images were embedded in the clean cover images dataset, which resulted in 10,000 2LSB stego images dataset and 10,000 4LSB stego images dataset.

4.2 Feature extraction

The selected features were extracted from the clean cover images and the two stego images datasets. The feature extraction program takes as input a batch of clean or stego image, and produces a feature vector for each image, which consists of 22 feature elements. The set of vectors for a batch of images are stored in a CSV file for subsequent classification. The GLCM and Entropy features are measured using the related functions in MATLAB.

The generated feature datasets of clean, 2LSB and 4LSB images have been made available online on [18].

4.3 Training / Testing

In this experiment we used the hold-our approach of train / test splitting ( ), in which 50% of the images are used for training and the other 50% for testing, to validate model accuracy. The extracted features datasets of the 10,000 clean images, 10,000 2LSB stego images and 10,000 4LSB stego images were split and combined to form the following training / testing datasets:

a. Training set 1 (TR1): consists of 5000 feature vectors of the first 5000 clean cover images, combined with 5000 feature vectors of the first 5000 2LSB stego images.

b. Training set 2 (TR2): consists of 5000 feature vectors of the first 5000 clean cover images, combined with 5000 feature vectors of the first 5000 4LSB stego images.

c. Validation set 1 (VS1): consists of 5,000 feature vectors of the second 5,000 clean cover images.

d. Validation set 2 (VS2): consists of 5,000 feature vectors of the second 5,000 2LSB stego images.

e. Validation set 3 (VS3): consists of 5,000 feature vectors of the second 5,000 4LSB stego images.

Class labels were attached to the training sets, comprising two categories, "clean" or "stego", whereas the validation sets were un-labeled, as they are "unseen records".

4.4 Classification

The selected classifier in this work is the Support Vector Machine (SVM) algorithm ( ) as implemented in MATLAB. The SVM classifier is chosen for its high accuracy as a two-category classifier, in addition to the high efficiency when dealing with large datasets.

To classify each validation set, the classifier was trained separately on the two training sets (TS1 and TS2), hence there were two classifications for the three validation sets (VS1, VS2, VS3). A feature vector, i.e. an image, is classified as stego if one or two of the classification runs produced stego output. The double classification is meant to provide blind steganalysis by training separately on 2LSB and 4LSB data, regardless of the validation data; which can be extended later to add other steganography schemes' training data.

4.5 Evaluation metrics

The following evaluation metrics are used in this work:

TP: True Positive rate, the ratio unseen stego images classified as such, to the total number of unseen stego images.

FP: False Positive rate, the ratio of unseen stego images classified as such, to the total number of unseen stego images.

TN: True Negative rate, the ratio of unseen clean images classified as such, to the total number of unseen clean images.

FN: False Negative rate, the ratio of unseen clean images classified as such, to the total number of unseen clean images.

Detection Accuracy: the ratio of true classification of unseen clean and stego images to the total number of unseen clean and stego images:

$$\text{Detection Accuracy} = \frac{TP + TN}{TP+FP+TN+FN} \tag{1}$$

5 RESULTS AND DISCUSSION

The three unseen clean and stego validation sets were classified using double classification. Each run classified a 5,000 unseen validation set twice, each time against training on a mixed 10,000 labeled training set. The average time for the dual training on 10,000 feature vectors and the classification of 5,000 feature vectors on each of the training sets, using MATLAB 2011b on a standard I3 PC, was under 50 seconds, which is an indication of the high efficiency of the SVM algorithm.

Table 2 shows the combined classification results of 5,000 clean images and 5,000 2LSB stego images, and Table 3 shows the combined classification results for 5,000 clean images and 5,000 4LSB stego images.

Table 2: Classification results of 5,000 clean images and 5,000 2LSB stego images

Evaluation Metric	Result
TP	99.22%
FN	0.78%
TN	98.82%
FP	1.18%
Detection Accuracy	99.02%

Table 3: Classification results of 5,000 clean images and 5,000 4LSB stego images

Evaluation Metric	Result
TP	100.00%
FN	0.00%
TN	98.82%
FP	1.18%
Detection Accuracy	99.41%

The results show a high detection accuracy of the clean and stego images, with the combined clean-4LSB stego having slightly higher detection accuracy at 99.41% due to the fact that the 4LSB stego scheme embedded twice the amount of the embedded data.

Also, the True Positive detection rate is slightly higher than the True Negative, which indicates that the model is inclined to do better in detecting stego images than clean images.

6 CONCLUSION

The work in this paper demonstrated that by focusing on the parts of a cover image where it is more likely to be used for embedding, the steganalysis can achieve high detection accuracy. The proposed feature set included statistical texture features of GLCM, Entropy of the right half-bytes, correlation between left and right half-bytes, coefficient of variation of right half-bytes, and the absolute difference between successive right half-bytes. The obtained results of analyzing 10,000 of each of clean, 2LSB stego and 4LSB stego images, using the SVM classifier have shown high detection accuracy (99.41% for 4LSB and 99.02% for 2LSB).

The present work can be extended for steganalysis under different situations, as in the suggestions below for future work:

- Extending and investigating the proposed feature set for the steganalysis of RGB images, with and without compression.
- Evaluating the proposed work with other steganography schemes.
- Investigating the effect of reducing the embedded data size on the detection accuracy and ways to avoid a decline in accuracy as a result of smaller embedded data size.

7 ACKNOWLEDGMENT

Our thanks to Middle East University, Jordan, for supporting this research work.

8 REFERENCES

[1] Schaathun, H. G. 2012. Machine Learning in Image Steganalysis, Wiley-IEEE Press, 1st edition.

[2] Kaur, M. and Kaur, G. 2014. Review of Various Steganalysis Techniques, *International Journal of Computer Science and Information Technologies*, Vol.5 (2).

[3] Pevný, T., Fredrich, J. and Ker, A. D. 2008. Novelty Detection in Blind Steganalysis, *Proc. ACM Multimedia and Security Workshop*, Oxford, UK, September 22-23, pp.167-176.

[4] Pevný, T., Fredrich, J. and Ker, A. D. 2011. From Blind to Quantitative Steganalysis, *IEEE Trans. on Info. Forensics and Security.*

[5] Liu, Q. and Chen, Q. 2014. Improved Approaches with Calibrated Neighboring joint density to Steganalysis and Seam-carved Forgery Detection in JPEG images", *ACM Transactions on Intelligent Systems and Technology*, 5(4), article 63.

[6] Aljarf, A., Amin, S., Fillipas, j. and Shuttelworth, J. 2016. The Development of an Image Detection System Based on the Extraction of Colour Gradient Co-occurrence Matrix (CGCM) Features, *Dese2016 Conference*, UK.

[7] Haralick R.M., Dinstein I, S. K. 1973. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics, 3: 610–621.

[8] Aljarf, A., Amin, S., Fillipas, j. and Shuttelworth, J. 2013. Develop a Detection System for Grey and Colour Stego Images, *International Journal of Modeling and Optimization, Vol. 3, No. 5.*

[9] Dalal, N. and Triggs, B. 2005. Histograms of Oriented Gradients for Human Detection, *9th European Conference on Computer Vision. San Diego, CA.*

[10] Mikolajczyk K, and Schmid, C. 2005. A Performance Evaluation of Local Descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29: 1615–1630.

[11] Tuzel O., Porikli, F., and Meer, P. 2006. Region Covariance: A Fast Descriptor for Detection and Classification. 9th European Conference on Computer Vision. 589–600.

[12] Ojala, T., Pietikainen, M, and Maeenpaa, T. 2002. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24: 971–987.

[13] Gong, R., and Wang, H. 2012. Steganalysis for GIF Images Based on Colors-Gradient Co-occurrence Matrix. *Optics Communications* 285: 4961–4965.

[14] Ker, A. D. 2004. Improved Detection of LSB Steganography in Grayscale Images, *International Workshop on Information Hiding*, IH 2004, pp 97-115

[15] Pevný, T. BOSSBASE1.01 Grayscale Database, DOI= http://agents.fel.cvut.cz/stegodata/PGMs/ downloaded on 2/1/2017.

[16] Franzen, R. Kodak Lossless True Color Image Suite, DOI= http://www.r0k.us/graphics/kodak/ downloaded on 2/1/2017

[17] Gonzalez, R. C. and Woods, R. E. 2017. Digital Image Processing, 4th ed., Prentice Hall

[18] Al-Jarrah, M. 2017. “Grayscale Steganalysis Features Dataset, Mendeley Data, v1. DOI= http://dx.doi.org/10.17632/w4fnr3r2j4.1#file-b47fc67e-60d6-49f0-a5c1-f7da1c8c289f