International Journal of Intelligent Computing and Information Science

A ROBUST STEGANALYSIS METHOD FOR DETECTING THE STEGANOGRAPHY IN IMAGES

M. N. Alkhalidi Osama Abu-Elnasr T. Elarif

Computer Science Department, Faculty of Computers and Information, Mansoura University - Egypt
al_khalidiii987@yahoo.com mr_abuelnasr@yahoo.com

Computer Science Department, Faculty of Computers and Information, Ain Shams University - Egypt
taha_elarif@yahoo.com

Abstract: *Recently, steganography and steganalysis have been received an increasing attention due the nature of our modern societies which depends on exchanging information on a large scale. Steganography is the art of communication through sharing secret messages by embedding them into useless cover messages. The cover message can be an image, audio, or video file. On the other side, the steganalysis techniques are concerned with discovering the existence of steganography. This paper presents a specific image steganalysis technique with main objective is to detect the existence of steganography made by the least significant bit (LSB) technique in a certain image. The proposed approach extracts the gray level co-occurrence matrix (GLCM) as salient features which capable to distinguish a stego image from a non-stego one using a Back-Propagation (BP) classifier at the classification phase. Experimental results on standard datasets that consists of 297 images are encouraging. The proposed method is robust and high accuracy level has been achieved.*

Keywords: *Steganography, Steganalysis, LSB, GLCM, BP.*

1. Introduction

The emergence of computers and internet has affected our societies in a great manner. The huge advancements in these fields have led to what we call the revolution of information. As a consequent of this revolution, there exist tremendous amount of information that need to be exchanged everyday securely. In order to achieve the information security, a set of techniques have been developed such as cryptography and steganography. The subtle difference between cryptography and steganography is that the aim of cryptography is to make the content of the message looks like rubbish while the aim of steganography is to hide the existence of the message itself [1]. Therefore, steganography is more effective than cryptography in achieving the information security. Steganography means hiding information in a certain way such that it cannot be detected in a cover media file. The cover media can be video, voice, image, etc. [2]. Unfortunately, the steganography can be misused by malicious hackers and intruders for passing a message which can cause catastrophic situations. So, sometimes, it is necessary to detect the stego file to prevent such situations. But because of the huge amount of information, it is impossible to achieve that manually.

Therefore, there was a necessity to develop techniques which capable to discover the existence of steganography. Later, these techniques called steganalysis techniques [3].

On the contrary, the steganalysis is the science which concerned with the detection of the existence of steganography in a certain media file [4]. In this paper, we are interested in a special kind of steganalysis techniques where the cover media file is an image. This type is called image steganalysis. Generally, image steganalysis can be categorized into two types: specific and generic. Specific steganalysis approach is being developed to detect the steganography made by a certain steganography technique. On the other side, the generic steganalysis approach is being developed to detect the hidden information regardless the used steganography technique [5]. Also, Steganalysis methods can be classified into two types according to the method of detecting hidden messages: Statistical steganalysis (in spatial domain or transform domain) and feature based steganalysis [6]. In this paper, we have proposed a specific image steganalysis technique with main objective is to detect the existence of steganography made by the least significant bit (LSB) technique in a certain image. The main phases of the proposed system are: feature extraction, feature reduction, classification. The details of each phase are introduced later in the following section.

The rest of this paper is organized as follows: section 2 explores some previous efforts. After that, Section 3 provides the necessary background concepts to understand the proposed work. Then, the proposed system is introduced in Section 4 followed by the experimental results in section 5. Finally, the paper is concluded with the conclusion and future work in section 6.

2. Related Work

During the last decade, steganalysis has been received a lot of attention from many researchers. This section covers some of the efforts which have been done in this important research area. In [7], a steganalysis technique has been proposed image steganalysis. The salient extracted features are the GLCM matrix in spatial domain. They considered many combinations of the diagonal elements of gray level co-occurrence matrix as features that used this feature to distinguish between stego and non stego image. For classification, they have used the Euclidean distance and Absolute distance to make their decision. Also, in [8], they have proposed a new method for detection hidden data for image setganalysis based on using neural network technique , that used Redial Based Neural Network (RBNN) for distinguishing normal from stego image. They extracted statistical feature from Karhunen-Loève (KL) transform coefficient obtained from co-occurrence matrix. Another steganalysis technique has been proposed in [9] called Visual Pixel Detection (VPD). Their method used for image steganalysis. The experimental results have shown that the proposed work gives a better performance than many well-known steganalysis techniques.

In [10], they have proposed a steganalysis technique which based on Spatial Gray Level Dependence Method (SGLDM) for texture analysis. For the classification process, the proposed method employed a back propagation (BP) neural network which trained using the texture related statistics extracted to discriminate between the stego and normal images. Also, the researchers in [11] have proposed a new method for steganalysis of gray image. Their proposed method used GLCM in addition to high order statistics in Discrete Wavelet Transform (DWT) coefficient as salient features. In the classification stage, they have used a Support Vector Machine (SVM) classifier. In [12], they proposed a steganalysis technique to detect the steganography made by LSB steganography algorithm. They used GLCM as distinguishing features. For the classification process, they employed a Multi-Layer Perceptron (MLP) neural network. In

[13], they have proposed method for detection hidden data based on feature extraction from Markov, histogram and Co-occurrence from wavelet domain and compared with existing farid 72 DWT features. For hide data in image it was used tow algorithms nsF5 and outguess. At last used BP classifier for distinguish between stego and normal image. Finally, in [14], they have proposed method for stego analysis based on feature extraction as Markov features, calibrated, and combined features of modified Discrete Cosine Transform (DCT). Then, they used (SVM) and (MLP) classifiers for image classification into two groups (normal and stego).

3. Background

This section provides the necessary knowledge and explains the needed concepts required to understand the proposed method. Starting with the GLCM features followed by the feature reduction method PCA then concluded with a brief description for different classification methods which have been used in the proposed work.

3.1 Grey-Level Co-occurrence Matrix (GLCM) Features

The GLCM is a robust way in statistical images analysis. It's used to evaluation of images features regarding to second-order statistics, by looking at the link between two neighboring pixels in one offset as the second order texture. By default, the GLCM is obtained by computing how many times a pixel with a gray level i exist horizontally adjacent to a pixel with a gray level j. The first pixel is named reference pixel while the second pixel is named neighbor pixel. However, we can determine other pixel spatial relationships by using different values for the 'offset' parameter. The default case is obtained by offset value equal to [0 1] where the first value represents the number of rows between the reference pixel and the neighbor pixel while the second value represents the number of columns between the reference pixel and the neighbor pixel. GLCM is usually defined as a two-dimensional matrix of joint probabilities between pairs of pixels [15]. The co-occurrence matrix is a statistical model that benefits different application in the images analysis, as in biomedical, etc. [16]. Many different statistical features (18 features) can be extracted from the GLCM such that energy, entropy, variance, etc. These features are computed based on a group of second order statistics [11]. The different feature names and equation are shown in table 1.

Table 1. The names and equations of different features extracted from the GLCM [11].

No.	Feature Name	Equation
1	Energy	$f_1 = \sum_i \sum_j \{P(i,j)\}^2$
2	Contrast	$f_2 = \sum_{n=0}^{n-1} n^2 \{ \sum_{i=1}^{N} \sum_{j=1}^{N} p(i,j) \ \|i-j\| = n \ \}$
3	Correlation	$f_3 = \frac{\sum_i \sum_j (ij)\, p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$

4	Entropy	$f_4 = -\sum_i \sum_j p(i,j) \log(p(i,j))$
5	Homogeneity	$f_5 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$
6	Autocorrelation	$f_6 = \sum_i \sum_j (ij)\, p(i,j)$
7	Dissimilarity	$f_7 = \sum_i \sum_j \lvert i-j \rvert \cdot p(i,j)$
8	Cluster Shade	$f_8 = \sum_i \sum_j (i-j-\mu_x-\mu_y)^3 \cdot p(i,j)$
9	Cluster Prominence	$f_9 = \sum_i \sum_j (i+j-\mu_x-\mu_y)^4 \cdot p(i,j)$
10	Maximum Probability	$f_{10} = \max_{i,j} p(i,j)$
11	Variance Sum Of Squares	$f_{11} = \sum_i \sum_j (i-\mu)^2 \cdot p(i,j)$
12	Inverse Difference Moment	$f_{12} = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$
13	Sum Average	$f_{13} = \sum_{i-2}^{2N} j\, p_{x+y}(i)$
14	Sum Variance	$f_{14} = \sum_{i-2}^{2N} (i-f_s)^2 p_{x+y}(i)$
15	Sum Entropy	$f_{15} = -\sum_{i=2}^{2N} p_{x+y}(i) \log\{p_{x+y}(i)\}$
16	Difference Variance	$f_{16} = Variance\ of\ p_{x-y}$
17	Difference Entropy	$f_{17} = -\sum_{i=0}^{N-1} P_{x-y}(i) \log\{P_{x-y}(i)\}$

18	Information Measure Of Correlation(1) Information Measure Of Correlation(2)	$f_{18} = \frac{HXY - HXY1}{\max\{HX, HY\}}$ $HXY = -\sum_i \sum_j p(i,j) \log(p(i,j))$ $HXY1 = -\sum_i \sum_j P(i,j) \log\{(P_x(i)P_y(j)\}$

3.2 Back-Propagation Neural Network

Back-propagation (BP) algorithm is a supervised training algorithm that is widely used by developers that working in artificial neural network. Error correction can be considered the essence of the learning algorithm. BP algorithm, mainly, consists of two phases: a forward phase and a backward phase [17]. The BP algorithm is usually be used for training multi-layer neural network which consists of one input layer, one or more hidden layer, and one output layer. We used (BP) algorithm for distinguishing normal from stego image.

4. The Proposed System

We have proposed a specific image steganalysis technique with main objective is to detect the existence of steganography made by the Least Significant Bit (LSB) technique in a certain image. A detailed block diagram of the proposed method is shown in Figure 1.

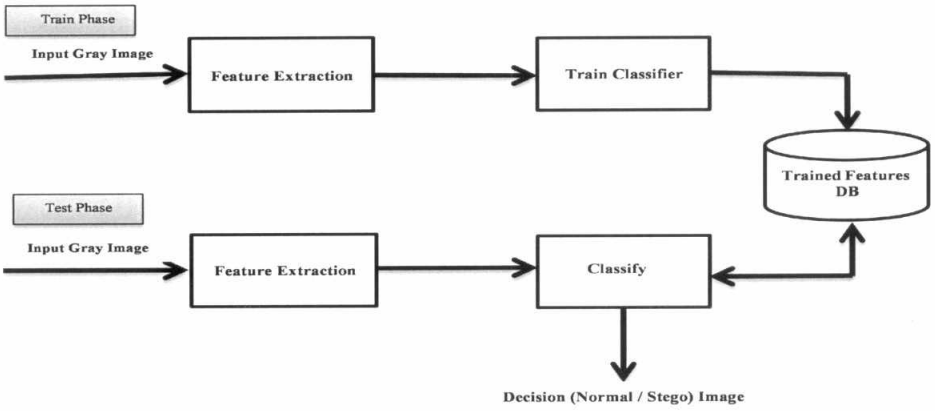

Figure 1. A detailed block diagram for the proposed method.

4.1 Feature Extraction

This stage is the most important one in our system because the accuracy of the classification process depends on the quality of the extracted features. It takes the input images (gray image) and returns the

GLCM. Then, the GLCM is analyzed by computing 18 statistical values which represent the extracted features. In our work, we have tried different offset values in computing the GLCM matrix. However, the best classification accuracy has been obtained with offset values [2 0] and [2 0, 0 2] which we will refer to as offset 1 and offset 2 respectively in the rest of the paper.

4.2 Training the Classifier

In this work, BP trained supervised classifiers were used. During training phase, a set of labeled examples (input, target) are used where the input are the feature or the reduced features while the target or the desired output is normal or stego image. The performance of the different classifiers is measured during a test phase as shown later in section 5.

5. Experimental Results and Analysis

A standard database called Berkeley Segmentation Data Set (BSDS) [18], has been used which consists of 297 color images with file extension .jpg. The data set is divided into 198 images for training and 99 images for testing. Initially, all of these images are clean or normal images. Then, these images are copied and turned into stego images using the LSB steganography algorithm. so, the total number of images for training becomes 396 images (198 normal and 198 stego) and the total number of images becomes 198 images (99 normal and 99 stego). The proposed system was implemented under a MATLAB2015 platform, windows7 OS, and Intel core i5-2400@1.6GHZ CPU.

Two different experiments have been conducted to evaluate the performance of the classifiers. The first one used offset1 features while the second one used offset 2 features. Results are shown in table 2.

Table 2. The results of the proposed system using BP classifier based on offset 1 and offset 2.

The sets	No. of images in the tests	Offset 1		Offset 2	
		Ratio of correctly classified images	Ratio of incorrectly classified images	Ratio of correctly classified images	Ratio of incorrectly classified images
Clean images	99	94.94 %	5.05 %	96.9 %	3.03 %
Stego images	99	98.98 %	1.01 %	100 %	0 %
Total	198	96.96 %	3.03 %	98.4 %	1.51 %

From Table 2 we conclude that the Offset 2 get better result than the Offset 1.Where the accuracy of Offset 2 is 98.4 %, but the accuracy of Offset 1 is 96.96 %. Where the ratio of correctly classified image from clean image increased from 94.94 % in Offset1 to 96.9 % in Offset 2, and the ratio of correctly classified

image from stego image increased from 98.98% in Offset1 to 100% in Offset 2. While the ratio of incorrectly classified image from clean image decreased from 5.05 % in Offset1 to 3.03 % in Offset 2, and the ratio of incorrectly classified image from stego image decreased from 1.01 % in Offset1 to 0 % in Offset 2. We note the Offset 2 give high ability to the BP for detect stego or clean image.

Finally, Table 3 shows comparisons between the proposed methods and the methods, which have been proposed in [10], and [19].

Table 3. The accuracy of different steganalysis methods.

No.	Classifier	No. of Features	Accuracy
[10][2013]	BP	5	82.88 %
[19][2014]	BP	6	95 %
Proposed Approach	BP	18	98.4 %

From Table 3 , We note the number of features that used in [10] and [19] is 5,6 features respectively, but we used in the proposed approach 18 features , Resulting the accuracy is increased to 98.4 % and that gives a high ability to the proposed approach to distinguish between stego or clean image.

6. Conclusion

This paper introduces a specific image steganalysis technique with main objective is to detect the existence of steganography made by the least significant bit (LSB) technique in a certain image. The proposed method extracts eighteen of the gray level co-occurrence matrix (GLCM) features which capable to distinguish a stego image from a non-stego one. The classification phase, BP classifier has been used. Experimental results show that the proposed approach give better accuracy results when compared with other work.

References

1. Turabieh, Hamza, et al. "Steganalysis of LSB encoding in digital images using GLCM and Neura Networks". (2006).
2. Sunaina Verma , Sandeep Sood, Sukhjeet Kaur Ranade. "Relevance of Steganalysis using DIH on LSE Stegnography". International Journal of Advanced Research in Computer Science and Software Engineering, V 4, No. 2,pp 835-838, (2014).
3. Geetha, S., and N. Kamaraj. "Optimized image steganalysis through feature selection using MBEGA" International Journal of Computer Networks & Communications, V.2, No.4,pp 161-175, (2010).
4. Kaur, Manveer, and Gagandeep Kaur. "Review of Various Steganalysis Techniques". International Journa of Computer Science and Information Technologies V.5, No.2,pp 1744-1747, (2014).
5. Meghanathan, Natarajan, and Lopamudra Nayak. "Steganalysis algorithms for detecting the hidder information in image, audio and video cover media". International journal of Network Security & Its

application (IJNSA) 2.1 ,pp 43-55,(2010).

6. Badr, Sherif M., Goada Ismaial, and Ashgan H. Khalil. "A Review on Steganalysis Techniques: From Image Format Point of View". International Journal of Computer Applications 102.4 (2014).
7. Kekre, H. B., A. A. Athawale, and Sayli Anand Patki. "Steganalysis of LSB embedded images using gray level co-occurrence matrix". International Journal of Image Processing (IJIP) v5,No.1, (2011).
8. Safwan O. Hasson Farhad M. Khalifa." Steganalysis Using KL Transform and Radial Basis Neural Network". Raf. J. of Comp. & Math's. , V. 9, No. 1, (2012).
9. Mansour, Romany F., W. F. Awwad, and A. A. Mohammed. "A robust method to detect hidden data from digital images". Journal of Information Security V.3,No.2 (2012).
10. Nissar, Arooj, and A. H. Mir. "Texture Based Steganalysis of Grayscale Images Using Neural Network" Signal Processing Research v.2.No.1,pp17-24 (2013).
11. Fazli, Saeid, and Maryam Zolfaghari-Nejad. "A new steganalysis method for steganographic images or DWT domain". International Journal of Science and Engineering Investigations 1, pp1-4, (2012).
12 Ghanbari, Sedighe, et al. "New steganalysis method using GLCM and neural network". International Journal of Computer Applications, V42, No.7,pp 45-50,(2012).
13. Saini, Manisha, and Rita Chhikara. "DWT Feature based Blind Image Steganalysis using Neural Network Classifier". International Journal of Engineering Research and Technology. V. 4. No. 04, (2015).
14. J. Anita Christaline, R. Ramesh and D. Vaishali. "steganalysis with classifier combinations". ARPN Journal of Engineering and Applied Sciences. V. 9, NO. 12,pp 2858-2863, (2014).
15. Benco, Miroslav, et al. "An Advanced Approach to Extraction of Colour Texture Features Based on GLCM". International Journal of Advanced Robotic Systems 11, pp 1-8, (2014).
16. M.Harsha vardhan, S.Visweswara Rao "GLCM ARCHITECTURE FOR IMAGE EXTRACTION" International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) V. 3, No. 1, pp.75-82, (2014).
17. Anitha, P. T., M. Rajaram, and S. N. Sivanandham. "neural network based steganalysis framework to detect stego-content in corporate emails". International Journal of Emerging Technology and Advanced Engineering Website, V.2, No. 3,pp 418-424,(2012).
18. https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/fg (online):[access at 5/2/2016].
19. Anjani Kumar Verma." A Non- Blind Steganalysis Through Neural Network Approach". International Journal of Multidisciplinary Consortium. V.1, No.1, pp 1-13,(2014).