Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

An Improved Approach to Blind Image Steganalysis Using an Overlapping Blocks Idea

*V. Sabeti**

Department of Computer Engineering, Faculty of Engineering, Alzahra University, Tehran, Iran.

Article Info

Article History:
Received 12 August 2022
Reviewed 18 October 2022
Revised 28 October 2022
Accepted 31 October 2022

Keywords:
Steganalysis
Steganography
Blind steganalysis
Block-based steganalysis
Spatial steganography
JPEG steganography

*Corresponding Author's Email Address: *v.sabeti@alzahra.ac.ir*

Abstract

Background and Objectives: Steganalysis is the study of detecting messages hidden using steganography. Most steganalysis techniques, known as blind steganalysis, focus on extracting and classifying various statistical features from images. Consequently, researchers continually seek to improve the accuracy of blind detection methods. The current study proposes a blind steganalysis technique based on overlapping blocks.

Methods: The proposed method began by decomposing the image into identically sized overlapping blocks, then extracted a feature vector from each block. Subsequently, a tree-structured hierarchical clustering technique was used to partition blocks into multiple classes based on extracted features, and a classifier was trained for each class to determine whether a block is from a cover or stego image. The block decomposition process was repeated for each test image, and a classifier was selected based on the block class to make a decision for each block. Furthermore, the majority vote rule was utilized to determine whether the test image is a cover or stego image.

Results: The proposed method was evaluated using the INRIA and BOSSbase datasets. Several parameters, including the number of block classes, feature extraction method, block size, and number and block overlapping level, affected the performance of the proposed method. The optimal block size was 64 × 64 by 32 steps, and the number of block classes was set to 16. WOW, S-UNIWARD, PQ, and nsF5 were the steganographic methods employed to evaluate the proposed method. Experimental results indicated that using overlapping instead of non-overlapping blocks increased the detection of data embedded in both the spatial and Joint Photographic Experts Group (JPEG) domains by an average of over 9%. In addition, the proposed method's accuracy in detecting the S-UNIWARD method was comparable to that of other deep learning-based steganalysis techniques.

Conclusion: The concept of using overlapping blocks improves the efficiency of blind steganalysis by providing the benefit of additional and larger blocks. One of the main advantages of the proposed method is comparable detection accuracy and less computational complexity than recent deep learning-based steganalysis techniques.

Introduction

With the advent of the web and its increasing use as a platform for digital data transmission, data protection techniques are more crucial than ever. To this end, numerous solutions have been developed for data security and safety. Individuals other than the sender and receiver are unable to understand the communication content due to data encryption before the transmission occurs through the network. Only the sender and receiver have access to the encryption key. Although the existence of encrypted communication is not hidden, steganography is the process of transmitting secret information by embedding it within a cover media. Consequently, this method conceals the existence of confidential information [1].

The secret data and cover media used in steganography may be text, image, video, or audio files. Among the various techniques, image steganography (i.e., hiding secret data in an image) is the most popular and widespread because it facilitates the transfer of large volumes of images over the internet. Digital images and videos contain a high proportion of repetitive bits, making them more suitable for data hiding [2].

In image steganography, the original image used to carry secret information is known as a cover image. The image resulting from the embedding process is known as a stego image. The success of steganography is predicated primarily on the secrecy of the embedded concealed data. Image steganography comprises three main requirements: hiding capacity, imperceptibility, and security [3].

The embedding process used to hide secret data in the cover image forms the basis of the steganographic process. Since it is possible to embed secret data in the spatial and transform domains of the cover image, existing steganographic techniques can be classified based on the cover domain employed. Due to the simplicity of embedding and extraction operations, spatial domain embedding techniques are more popular and utilized than transform domain techniques. However, they possess less robustness and reliability. Embedding in Discrete Cosine Transform (DCT) coefficients of Joint Photographic Experts Group (JPEG) images is one of the transform domain techniques [4].

Steganalysis is the opposite of steganography, the art, and science of deciphering covert communications through in-depth knowledge of steganographic techniques. Steganalysis is the science of attacking steganography in a never-ending battle with the primary objective of gathering sufficient evidence of the presence of an embedded secret message and breaching the security of the message carrier. Steganalysis methods are generally classified into three visual, structural, or statistical categories [5]. Visual steganalysis approaches, the simplest form of steganalysis, seek to detect visual anomalies within the stego image. Numerous visual steganalysis methods rely on deficiencies in embedding algorithms. Structural steganalysis detects modifications to the stego file format and reveals the presence of embedded data by comparing the structure to its standard set. As the most prevalent available techniques, statistical steganalysis techniques uncover covert data by comparing the statistical characteristics of a stego image to a set of covers.

Alternatively, steganalysis techniques can be classified as either special or universal steganalysis [6]. Special steganalysis methods are designed for a particular steganographic algorithm. In contrast, universal or blind steganalysis is a general technique that can decipher data embedded by any steganographic algorithm, even a previously unknown one.

Recent steganalysis literature has primarily focused on blind statistical steganalysis methods (referred to as blind steganalysis in the present study). Enhancing the quality of extracted feature vectors from images is one strategy for improving the performance of blind steganalysis algorithms. The richer this vector is with informative features, the better the algorithm's performance. Therefore, the quantity and quality of image features extracted have become crucial for blind steganalysis design. Indeed, several recent studies also employed deep learning algorithms, such as Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and other algorithms in which feature extraction and selection are performed automatically [7].

Frame-based steganalysis and block-based steganalysis are two blind steganalysis approaches that use the entire image or image blocks to extract features, respectively. The complete structure of block-based steganalysis methods is presented in [8]-[10]. The present study uses image blocks instead of the whole image in the feature extraction process. The majority of existing methods employ completely distinct and non-overlapping blocks. In addition to block images, this paper proposes the concept of overlapping image blocks, which significantly improves detectability by coordinating the number and size of image blocks. The implementation results demonstrate that the proposed method significantly outperforms its predecessors. The following summarizes the study's main contributions:

- Provide block-based steganalysis methods that use overlapping blocks for feature extraction.
- Investigate the influence of parameters, including the number of block classes, the feature extraction method, the size and number of blocks, and the degree of overlapping blocks, on the accuracy of the proposed approach.
- Evaluate the probability of using the proposed method to discover steganographic techniques in the spatial and JPEG domains.

The remainder of the paper is organized as follows: Section 2 presents several block-based steganalysis techniques after introducing the structure of blind statistical steganalysis methods. Section 3 describes the overlapping blocks-based steganalysis method. Section 4 presents the results of implementing and applying the proposed steganalysis approach in detecting several existing JPEG and spatial steganographic methods. Finally, Section 5 provides conclusions and recommendations for future research.

Related Work

Steganalysis is said to be successful only if the hidden message embedded in media is proven. Recent steganographic techniques attempt to leave cover media with minimal quantitative and statistical traces. Conversely, in response to this practice, standard steganalysis approaches attempt to broaden their analysis dimensions and employ complex and expert processes to achieve greater sensitivity. Therefore, modern steganalyzers require significantly more computing resources and power than in the past.

New approaches are required to conserve resources and simplify steganalysis, reducing computational complexity and time while increasing productivity [12]. The blind statistical steganalysis methods are comparable to pattern classification techniques. After applying some image preprocessing operations, most existing blind steganalysis methods extract a vector of features from images. Then, they select or design a suitable classifier and train it using the extracted features from the training image set. The training images consist of both cover and stego images. The output of the training phase is a classifier that can be used to determine the state of the test images. After applying image preprocessing operations and feature vector extraction in the testing phase, the trained classifier classifies the test image into one of two cover or stego categories. Blind steganalysis general steps are depicted in Fig. 1 [11].

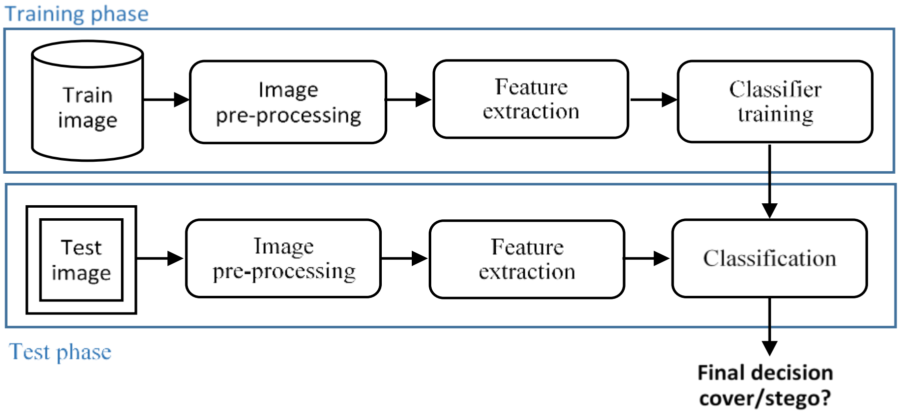

Fig. 1: The framework of blind steganalysis methods [11].

In the image preprocessing step, some operations are applied to the image before feature extraction, including converting RGB images into grayscales, cropping, JPEG compression, and DCT transformation, among others. The feature extraction step includes a set of unique statistical properties of an image, referred to as features. This step attempts to extract informative image features. The informative feature indicates that the selected feature must be embedding sensitive information. Following feature extraction, a low-dimensional feature vector is constructed, which reduces the computational complexity of the training and classification phases. Feature extraction techniques are wavelet decomposition, Markov empirical transition matrix, image quality metrics, co-occurrence matrix, and image histogram, among others [13]. Currently, existing steganalysis methods, such as special and blind methods, aim to enhance detectability and efficiency [7]. In this regard, numerous advancements have been made by researchers. Some extract more features from existing images [6], whereas others search for the optimal, highest -quality, and smallest feature vector without increasing the number of features [12]. Another group of methods selects the desired features from the specific parts of the images [14], while others focus on improving the classifiers' performance [15].

Since the performance of blind steganalysis algorithms is highly dependent on the quality of the extracted features, a line of related research was devoted to determining the requirements for feature extraction. Several of these studies attempted to identify the most probable location for information embedding and feature extraction based on the type of cover image, image texture, and image color, among other factors [16].

Images are comprised of different decompositions with varying complexities and frequencies. When hidden information is embedded into an image, the effect is more noticeable when the image is less complex. A single image may also contain heterogeneous regions, and some of its decompositions may have greater complexity and frequency rates. Consequently, block-based steganalysis methods emerged in response to these challenges and observations [8]. The central idea is to decompose images of comparable complexity into smaller, more uniform blocks than the original image. Then, each obtained block is considered a discrete steganalysis input image. It has been demonstrated that steganographic embedding correlates much more strongly with similar blocks. Thus, the features of these smaller blocks are used to create a content-based classifier. Table 1 summarizes several blind steganalysis methods that employ image blocking prior to feature extraction. These methods are described in greater detail below.

Table 1: The summery of block-based steganalysis methods

Ref.	Blocking	Overlapping	Texture analysis	Feature extraction method	Domain	Tested methods
[8], [9]	✓	×	×	Pevny's method	Transform (JPEG)	OutGuess, F5, MBS
[10]	✓	✓	×	SPAM	Transform (JPEG)	PQ, MBS
[17]	✓	×	✓	Merged-274 feature set	Transform (JPEG)	F5, nsF5 MB1, PQ, JPHide
[18], [19]	✓	×	×	IQM	Transform (JPEG)	F5, Quickstego, StegHide
[14]	✓	×	✓	SPAM CC-PEV	Spatial	HUGO
[20]	✓	×	✓	SRM	Spatial	LSBR, LSBM, HUGO, S-UNIWARD

Studies [8]-[10] can be cited as being among the first to utilize the idea of block-based steganalysis. The image is first decomposed in these methods into smaller, fixed-size blocks. Based on the feature vector of each block, the blocks are then decomposed into multiple classes. For each block class, a classifier is trained based on the features extracted from each block in that class. Similar to the training phase, the block and feature extraction operations are repeated for each input image during the testing phase. Then, each image block is assigned to a particular class, and the classifier for each class determines whether the block is cover or stego. The final step is to combine the results of all classifiers regarding the type of blocks to determine whether the entire image is a cover or stego.

Wang et al. [17] proposed a block texture-based cover image method for JPEG image feature extraction and steganalysis. Their method decomposes the input images into several sub-images based on the JPEG block texture complexity. The calibrated set of features is then extracted from each of these subimages. Separate sets of subimages with the same texture complexity are used to construct and train the classifier. The end result of steganalysis is attained through a process of weighted fusing. Due to the insufficiency and limitation of the obtained feature set, this method lacks the optimal detection accuracy for detecting embedded images with a low rate or some novel and unknown methods.

Another disadvantage of this method is the significant computational complexity and computation time complexity of detecting image texture.

Suryawanshi et al. [18], [19] also presented a blind statistical method for digital image steganalysis. In this method, the image is first decomposed into identical blocks. Then, the statistical features of each block are extracted. Several sub-classifiers of a multi-class classifier are used to classify the image based on these features. The proposed scheme outperforms several existing approaches due to each block's initial image block and multidimensional feature extraction.

Mohammadi et al. [14] proposed a universal statistical steganalysis method that decomposes test images into sub-images using the Artificial Bee Colony (ABC) algorithm. Then, the optimal sub-region concerning density and energy is selected, and the desired features are extracted from this region. These two feature sets are combined to train the Support Vector Machine (SVM) classifier. Experimental results from the algorithm implementation demonstrate that the proposed method improves detection accuracy and increases True Positives (TPs) and True Negatives (TNs).

Zhu et al. [20] used image decomposition to introduce a block-based steganalysis method. This method decomposes the image into subimages with differing texture levels. Subimages are utilized for training the classifier, which aids in simulating statistical detectability.

This technique is only employed to decipher spatial steganographic techniques.

The Proposed Model

This paper proposes a block-based steganalysis framework. As mentioned earlier, other block-based steganalysis methods have also been presented. However, the advantage of the proposed method over other existing methods is that the overlapping blocks idea is used for a fixed-size image in addition to the number of blocks with specific dimensions. Therefore, since the number of blocks is not necessarily fixed in this technique, the number and size of the blocks can be coordinated appropriately. As the size of blocks increases, their number does not decrease, and the feature vector is rich enough to train the classifier used well. Fig. 2 illustrates the block diagram of the proposed block-based steganalysis system. This system includes training and testing phases, each of which will be discussed in detail in the subsequent sections.

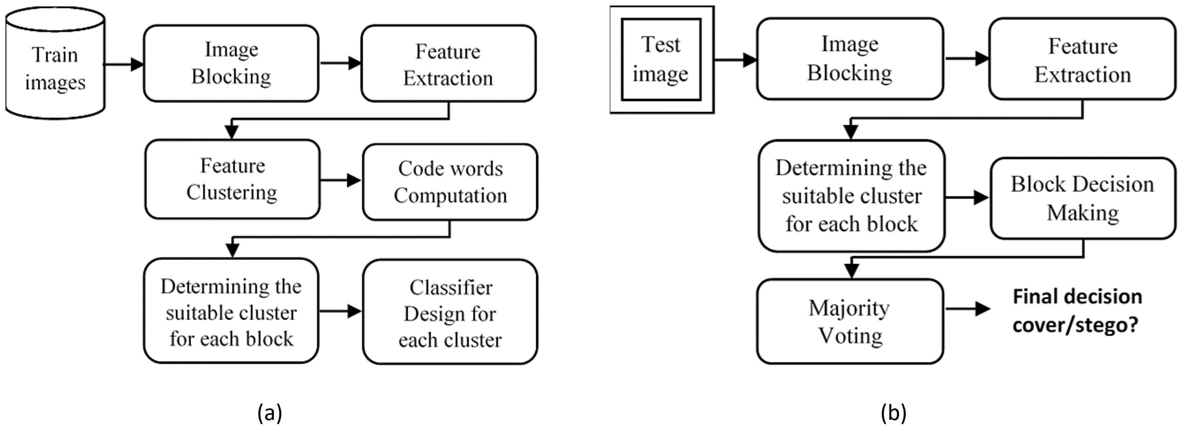

Fig. 2: The block diagram of the block-based steganalysis system (a) training phase (b) testing phase.

A. Training Phase

A set of images is required for the training process, including a combination of cover and stego images. After selecting a set of cover images, one or more steganographic algorithms are chosen to generate stego images. Then, using the steganographic algorithm, the secret data is embedded in the existing cover images to generate the corresponding stego image. This process creates the required set of cover and stego images. All the $M \times N$ cover/stego image pairs in the training set are decomposed into smaller homogeneous blocks of size $B \times B$. Then, the selected feature extraction algorithm is applied to all obtained blocks.

Block size and number affect the steganalysis's accuracy. There are some noteworthy points [10]:

- If the block size is enlarged, the standard deviation of the block features will decrease, and it will be easier to design a classifier that can distinguish between cover images and stego images. Therefore, the larger block size increases the extracted features' detectability.
- The more the number of blocks, the more accurate the results obtained from the classification stage.

Given these points, larger block sizes and more block numbers are more suitable for block-based steganalysis. However, there is an inverse relationship between the block size and the number of blocks in non-overlapping mode, and it is impossible to increase both parameters simultaneously. Using overlapping blocks is an alternative solution to achieve this goal. In other words, if overlapping blocks are used, the number of blocks will increase, assuming the use of fixed-size blocks. Nevertheless, it is noteworthy that the overlapping level of blocks is an influential parameter that can be changed to control the increase in the number of blocks. A trade-off between the detection accuracy and the number of blocks is required, as decomposing more blocks from the training set increase the computational complexity. If the number of decomposed blocks is vast, a random sampling algorithm is used to select a subset of blocks, which reduces the classification complexity.

Up to this point, the training set contains cover images and corresponding stego images. For K sampled blocks selected by random sampling algorithm, $K/2$ sampled blocks are chosen from cover images. The remaining $K/2$ blocks are corresponding blocks decomposed from stego images. Generally, random sampling is better than linear sampling since it selects more diverse blocks.

Then, we should look for a way to partition the blocks into different C classes. This partitioning is based on the steganalysis features decomposed from the blocks. Therefore, all blocks are clustered into different classes based on the similarity of their features. After clustering the blocks into different classes, the averaged feature vector is calculated in each block class, referred to as the codeword of that block class. For example, if we use merged Markov features [21] for feature extraction, each codeword will have 274 feature components. The more the number of C classes, the better the steganalysis performance, but the complexity is also increased. Therefore, we should seek a suitable C that balances computational complexity and performance.

The Tree-Structured Vector Quantization (TSVQ) method is used for block clustering, which converts image blocks into a binary tree structure based on similarity. Hence, the entire set of sampled blocks is divided into two subsets. This process is repeated in each subset until all blocks have similar features at a certain level. The K-means clustering algorithm is used in each partitioning step to decompose cluster S into two sub-clusters, which are denoted by S_1 and S_2. It is done by minimizing the within-cluster energy sum, i.e., $E(S_1.S_2)$, which is given by (1):

$$E(S_1.S_2) = \sum_{X_i \in S_1} \|X_i - \mu_1\|^2 + \sum_{X_i \in S_2} \|X_i - \mu_2\|^2 \qquad (1)$$

where $X_1.X_2 \dots X_n$ represent the feature vectors of n blocks, and μ_i denotes the averaged feature vector in S_i, which can be calculated according to (2):

$$\mu_1 = \sum_{X_i \in S_1} X_i \;.\; \mu_2 = \sum_{X_i \in S_2} X_i \qquad (2)$$

After dividing K blocks into different C classes, the averaged feature vector or codeword is determined for each class. Codewords are utilized to classify the blocks of a test image using the minimum distortion energy criterion in the feature space. In the implemented TSVQ, the partitioning operation stops when all the blocks within a node are homogeneous enough. Therefore, the stopping criterion in the clustering is based on $E(S_1.S_2)$ value, and the node with the most significant minimum distortion value is always split. This process is repeated until reaching the desired number of C classes.

After obtaining the C codewords representing the C classes of the sampled blocks, all the sample blocks in the training set are assigned into one of the C classes. This assignment is based on the distortion criterion $E_i(f_c.f_s)$, which is defined as the sum of two energies from a codeword for class i, according to (3):

$$E_i(f_c.f_s) = E_i(f_c) + E_i(f_s) \qquad (3)$$

where f_c and f_s represent the feature vectors of a block from the cover and stego images, respectively, $E_i(f_c)$ indicates the energy between the number of features extracted from a cover image block, and $E_i(f_s)$ denotes the same value for the stego image block.

After calculating $E_i(f_c.f_s)$ for all C classes, if $E_i(f_c.f_s)$ has the minimum value among all $E_i(f_c.f_s)$ for $1 \le i \le C$, the block pair of the cover image and its corresponding stego image is assigned into class C_j. Using the features of cover and stego image blocks for each class, a special classifier, such as an SVM classifier, can be trained for each of the C classes.

B. Testing Phase

Each test image, as in the training images, is blocked according to block size and blocks' overlapping level. Each block of the test image is assigned into a class using the minimum distortion energy. Depending on the class of each block, the classifier obtained from the training process is applied here. a decision is made about whether each block is a cover or a stego block. Therefore, every test image's total number of decisions equals the block number. Then, a majority voting approach is utilized to decide whether a given image is a cover or a stego. If the number of cover blocks is more/less than the number of stego blocks, the image is identified as a cover/stego.

Results and Discussion

The block-based steganalysis framework was implemented in the MATLAB software environment. All tests were performed on a computer with a Corei7 CPU with four cores and 6GB memory capacity. Therefore, the algorithm was tested on different steganographic methods in both spatial and JPEG domains to study the performance of the proposed block-based steganalysis algorithm. In each mode, 500 images were selected as cover and stego images (1000 images in total) in the training and testing phases.

The well-known and widely used BOSSbase dataset was used for the spatial domain. Some of these sample images are shown in Fig. 3. The dataset contains a thousand grayscale images with a size of 512×512 pixels in PGM format. The WOW and S-UNIWARD algorithms were used to embed messages in spatial domain.

Fig. 3: Sample images from the BOSSbase dataset.

The INRIA Holidays dataset was used for JPEG images. The target dataset contained more than 1,400 color images under JPEG compression with medium quality. One of the advantages of this dataset is that its images include different subjects, textures, complexities, and sizes. These images are not special, and mostly they are universal images. The selected images are transformed into 512×512 grayscale images to be used in the proposed algorithm. Some of these sample images are shown in Fig. 4. Perturbed Quantization (PQ) and nsF5 algorithms are used to embed the secret messages. Notably, images are compressed once again in the PQ method.

The intended embedding data must undergo encryption and compression processes before the embedding process to remove any semantic relationships between data bits, reduce the data size as much as possible, and turn the data into a bit string with random data properties; therefore, the different tests use random data (produced in MATLAB) with different lengths. Fig. 5 shows the resulting stego image from the 0.4bpp embedding level with different steganography methods in Lena image. Cover and stego images cannot be distinguished from one another by human eye. Therefore, visually, the output image of these methods is similar to the cover image, and the presence of data in these images can only be discovered through statistical analysis.

Fig. 4: Sample images from the INRIA dataset (right side: original images, left side: grayscale).

(a) Cover (b) Stego: PQ (c) Stego: nsF5 (d) Stego: WOW (e) Stego: S-UNIWARD

Fig. 5. Cover and resulting stego images from different embedding algorithms (0.4bpp).

Several parameters affect the performance of the proposed method, including the number of block classes (C), feature extraction method, block size and number, and block overlapping level. In the following, by changing these parameters, the accuracy of the proposed method for detecting steganographic methods in both spatial and JPEG domains is measured for different embedding rates or different Bit Per Pixels (BPP). Two versions of the proposed method were used in the tests performed. In the first version or NOBS, the image is decomposed to non-overlapping blocks, but in the second version or OBS, the overlapping blocks are used for the image blocks. The frame-based steganalysis method is also referred to as FS.

A. The Performance Comparison of Frame-Based and Block-Based Steganalysis Techniques

The idea of block-based steganalysis is proposed to improve the accuracy of frame-based steganalysis. Before examining the influence of the parameters on the proposed method's performance, this subsection examines the detection accuracy of steganalysis in two without block (frame-based method) and with block (non-overlapping) modes in both spatial and JPEG domains to compare and create an overview of the obtained results. Table 2 reports the steganalysis performance for different embedding rates of JPEG domain methods. In the NOBS, the number of block classes is 16, and the block size is

64×64. This test uses [22] for feature extraction and SVM with Gaussian kernel as a classifier. The results show that the NOBS steganalysis outperforms the traditional steganalysis methods. In addition, according to Table 1, the PQ has lower detection accuracy than the nsF5 methods, i.e., the PQ is more secure against steganalysis.

Table 3 shows the detection accuracy obtained for different embedding rates of the spatial domain. In the experiments, the method [23] was used as the FS method and NOBS method used the SPAM feature [23] for feature extraction phase. The number of block classes is 16, and the size of each block is 64×64 for NOBS. According to the results, NOBS's performance improvement over the FS method is evident in all embedding rates. Furthermore, the S-UNIWARD algorithm is slightly more robust against the steganalysis attacks than the WOW algorithm.

Table 2: The detection accuracy of FS and NOBS techniques for JPEG domain

	PQ		nsF5	
BPP	**FS**	**NOBS**	**FS**	**NOBS**
0.05	51.3	**52.72**	52.46	**54.50**
0.1	52.16	**55.45**	54.85	**58.72**
0.2	53.33	**60.45**	59.43	**66.20**
0.3	54.20	**67.04**	62.38	**73.50**
0.4	55.34	**73.18**	70.50	**82.75**

The results of Tables 2 and 3 demonstrate that as the embedding rate of the image increases, the NOBS exhibits higher detection accuracy. The following subsections examine the impact of the available criteria on block-based steganalysis performance. Notably, each phase selects the most robust algorithm for the experiments. Thus, for the spatial domain, the implementation results are tested on stego images created by the S-UNIWARD and images created by the PQ embedding technique for the JPEG domain.

Table 3: The detection accuracy of FS and NOBS techniques for spatial domain

	WOW		S-UNIWARD	
BPP	**FS**	**NOBS**	**FS**	**NOBS**
0.05	52.80	**54.30**	53.20	**53.90**
0.1	55.25	**59.70**	55	**58.45**
0.2	57.40	**66.5**	56.5	**65.85**
0.3	61.75	**74.04**	59.23	**73.33**
0.4	66.72	**84.20**	64.40	**82.85**

B. The Influence of Block Class Numbers

After blocking the input images, the decomposed blocks are classified into different classes based on the feature extracted through a classifier. It is expected that with the increase in the number of block classes (C) and more available codewords, the average block detection accuracy and the final detection accuracy of the algorithm will improve because of increasing the similarity of the block features belonging to each class. Table 4 reports the experimental results of the effect of block class number on the PQ steganalysis for the embedding rate of 0.6. As expected, the detection accuracy increases from 71.50 to 82.56 as the number of block classes increases from 2 to 64. However, the algorithm's performance improvement becomes saturated when the number of classes reaches 32 and more. Obviously, as the number of classes increases, more costs must be paid for computations. Therefore, experiments usually consider the middle bound for the number of classes to achieve optimal performance and balance the number of classes and computational complexity.

Table 4: The detection accuracy of the NOBS for different block classes

Number	Accuracy
2	71.05
4	73.27
8	75.31
16	78.53
32	81.70
64	**82.56**

C. The Influence of the Feature Extraction Method

Each feature extraction method focuses on specific image dependencies and statistical data. Since there are different feature extraction methods, the performance of available steganalysis algorithms will also vary. Therefore, this subsection examines the effect of different feature extraction methods on the proposed steganalysis performance. This study used three basic and well-known techniques to extract the desired features to measure the NOBS's performance for the JPEG domain. Other steganalysis researchers have widely used these techniques. Table 5 shows the obtained results.

This experiment used 512×512 images with 64×64 block size and eight block classes. Data were embedded into images with embedding capacity from 0.05 to 0.4. As can be seen, for low embedding capacities, the [24] method slightly outperforms the two methods [21], [22]. However, as the embedding capacity increases, [22] has the highest detection accuracy among these three approaches. In any case, [21] is the least influential compared to other tested methods. By default, all the experiments applied [22] for the feature extraction phase to achieve the best performance and detectability.

Two different feature extraction methods were used to investigate the effect of the extracted feature set on the spatial domain steganalysis. The first method is the SPAM [23], which extracts 686 features based on second-order Markov features from each image. The second method is the spatial rich model (SRM) [25] , which extracts 34671 features from each image. Table 6 reports the obtained results.

Table 5: The NOBS detection accuracy for the effect of different feature extraction methods on the PQ approach

	Feature extraction method		
BPP	**[22]**	**[21]**	**[24]**
0.05	51.59	50.45	**52.04**
0.1	55.22	53.86	**55.90**
0.2	57.04	55.90	**58.40**
0.3	**65.22**	57.95	61.81
0.4	**72.27**	68.40	70.02

Table 6: The NOBS detection accuracy for the effect of different feature extraction methods on the S-UNIWARD

	Feature extraction method	
BPP	**SPAM**	**SRM**
0.05	50.80	**52.80**
0.1	55.20	**57.20**
0.2	62.50	**65.5**
0.3	71.35	**75.35**
0.4	80.18	**81.18**

According to Table 6, the SRM outperforms the SPAM due to the high feature dimensions and richer feature vector. On the other hand, the steganalysis execution procedure is time-consuming due to the high computational complexity, and the performance decreases compared to the expected level. The following subsections use the SPAM feature extraction method for the spatial domain to facilitate the experimental process.

D. The Influence of Block Size and Number

As mentioned earlier, as the block size increases, the average block decision accuracy increases and improves the overall algorithm's detectability. As the number of blocks increases, the performance also improves due to the availability of more features. However, assuming non-overlapping blocks, in practice, there is an inverse relationship between these two parameters. For a fixed size 512×512 image, increasing the block size will result in fewer blocks, negatively impacting performance. Table 7 presents the average detection accuracy of the NOBS for different embedding rates on the PQ algorithm in the JPEG domain for 512×512 images and 16 block classes.

As the block size increases, which reduces the number of blocks for four different block sizes, the algorithm performance also decreases. Alternatively, the smaller the size of the blocks, the higher the algorithm's computational complexity due to the increase in the number of blocks. Fig. 6 depicts these results. According to the results, for lower embedding capacities, the detection accuracy obtained is relatively the same and somewhat negligible. However, the performance degrades as the data embedded in the image and the block size increase.

It is also worth noting that the experiments considered the smallest block size as 32×32. For smaller sizes, it is possible to reverse the results and reduce the detection accuracy due to excessive computational complexity. Therefore, the minimum threshold for the block size will be 32×32. Both the block size and the number of blocks individually have a significant impact on block-based steganalysis performance. Hence, their relationship should be balanced to achieve optimal performance and benefit from both positive effects.

Table 7: The NOBS detection accuracy for different values of block size and number

	Block size			
BPP	**32×32**	**64×64**	**128×128**	**256×256**
0.05	**55.22**	52.72	50.22	49.70
0.1	**57.54**	55.45	52.50	50.54
0.2	**63.18**	60.45	55.45	53.22
0.3	**72.40**	67.04	62.27	56.50
0.4	**78.53**	73.18	60.77	57.95

E. The Influence of Overlapping Blocks

The previous subsection investigated the inverse relationship between block number and size and their impact on the proposed steganalysis performance. According to the previous results, there is a dependency between block number and size, and an average limit should be considered for optimal performance. The proposed method uses overlapping blocks to overcome this limitation. For a block with a given size, the blocks are moved by predetermined step sizes (S) to get more. Therefore, one can take advantage of both larger blocks and more blocks. It will significantly improve the algorithm performance. Table 8 lists the number of blocks for a 512×512 image with different block sizes (four modes) and different step sizes (three modes). Note that if the step size equals the block size (S=B), the blocks are not overlapping. As the step size decreases, the overlapping degree of blocks and thus the number of blocks increase.

Table 9 presents the detection accuracy of the proposed method for both domains for two modes of non-overlapping blocks and overlapping blocks with a step size equal to one-half of the block size (S = B/2). In the experiment, the block size is 64×64 in normal mode. Consequently, by implementing the overlapping approach and considering a step size of 32 pixels for 512×512 images, we will have a fixed number of 225 blocks for each image. According to the results, the detection accuracy is improved as the number of overlapping blocks increases.

Table 9 shows the detection accuracy improvement rate for the PQ algorithm related to the JPEG domain for two modes based on blocks with a fixed number and overlapping blocks. Considering the step size equal to one-half of the block size, the number of blocks also increases in the spatial domain, and as a result, the algorithm performance improves compared to the previous state. The results of the S-UNIWARD experiments are also shown in Table 9. On average, the block-based steganalysis detection accuracy using the idea of overlapping blocks improves by more than 9% for PQ detection and more than 9.32% for S-UNIWARD detection at various embedding rates.

Table 8: Number of blocks for a 512×512 image with different block sizes and step sizes

S: Step size, B: Block size			
block size	**S = B**	**S = B/2**	**S = B/4**
256×256	4	9	25
128×128	16	49	169
64×64	64	225	841
32×32	256	961	3721

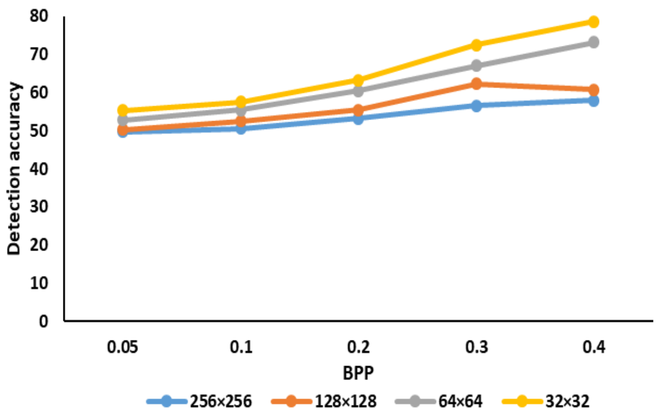

Fig. 6: Decreased NOBS detection accuracy with increasing block size.

Table 9: NOBS and OBS detection accuracy for two PQ and S-UNIWARD approaches at different embedding rates

	PQ			S-UNIWARD		
BPP	**NOBS**	**OBS**	**improvement**	**NOBS**	**OBS**	**improvement**
0.05	52.72	55.50	**5.3%**	53.90	56.20	**4.3%**
0.1	55.45	62.25	**12.3%**	58.45	66.70	**14.1%**
0.2	60.45	64	**5.9%**	65.85	74.25	**12.8%**
0.3	67.04	73	**8.9%**	73.33	79.5	**8.4%**
0.4	73.18	82.5	**12.8%**	82.82	88.64	**7%**

In general, for a given step size, the smaller the step size is, the more blocks are obtained, with fixed image size and fixed block size, and thus, the steganalysis performance is improved. However, despite the performance improvement, minimum step sizes are impractical due to the high computational complexity. By benefiting from overlapping blocks, the algorithm performance can be easily enhanced for different modes. For convenient comparison, Figs. 7 and 8 illustrate the bar charts of three steganalysis modes for the JPEG and spatial domains.

F. The Performance Comparison of the Proposed Method with State-Of-The-Art Techniques

The results of the earlier subsections demonstrate that if the steganalysis method uses the features extracted from the image to decide the state of the test image, the use of overlapping blocks (rather than the whole image or non-overlapping blocks) for feature extraction can enhance the accuracy of the steganalysis techniques. However, several novel steganalysis approaches use deep learning algorithms in which feature extraction is performed automatically. For better performance measurement, Table 10 presents the comparison results of the proposed approach with six deep learning-based steganalysis methods for S-UNIWARD at different embedding rates. These results were extracted from [26], [27] and [28], and in cases marked with a dash (–), the desired result was not cited in the references.

Table 10 reveals that the automatic feature extraction approaches are not significantly superior to the proposed method. Another main advantage of the proposed method is the much lower computational complexity. The statistics provided in [27] indicate that the time required to train the fastest method available [28] is about 3 hours. In contrast, less time is needed for non-automatic feature extraction-based steganalysis techniques. Indeed, despite the growing research of deep learning-based steganalysis methods and the existence of potent state-of-the-art hardware, it is not unlikely that these methods will be developed very quickly.

Table 10: The comparison of the OBS detection accuracy and deep learning-based methods at different embedding rates

BPP	Xu-Net [32]	Ye-Net [31]	SN-Net [30]	ReST-Net [29]	CIS-Net [28]	SFR-Net [27]	AG-Net [26]	OBS
0.1	59.43	59.17	64.79	64.15	64.72	-	-	**66.7**
0.2	66.67	66.49	73.18	68.73	73.79	**76.8**	-	74.25
0.3	73.68	74.38	79.29	76.44	76.36	-	**80.66**	79.5
0.4	80.12	77.36	83.47	84.28	85.38	87.9	85.49	**88.64**

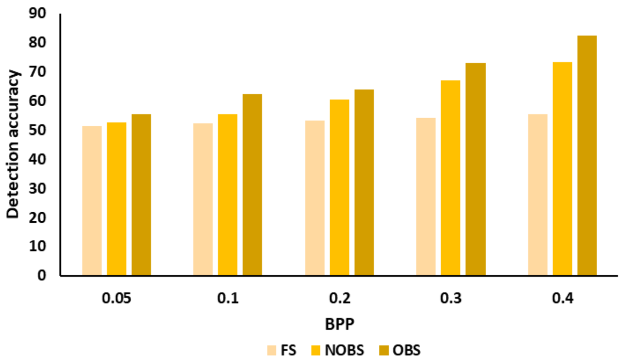

Fig. 7: The detection accuracy of FS, NOBS, and OBS approaches for PQ.

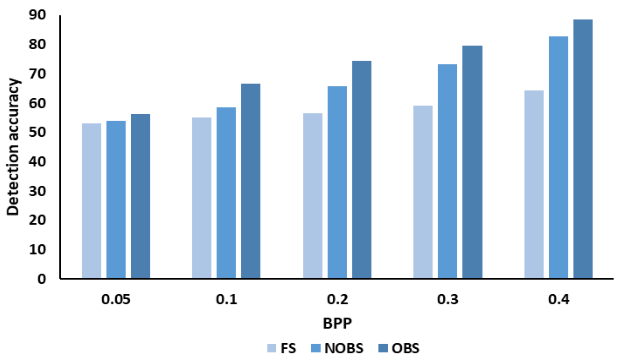

Fig. 8: The detection accuracy of FS, NOBS, and OBS approaches for S-UNIWARD.

Conclusion

Most steganalysis techniques require image feature extraction to detect stego images significantly altered by data embedding. Some steganalysis methods use the entire image to extract features (frame-based steganalysis), while others decompose the image into blocks and perform feature extraction for each block separately (block-based steganalysis). Block-based steganalysis has two major advantages over frame-based steganalysis. (1) the accuracy of block-based steganalysis is improved without increasing the number of features, and (2) block-based steganalysis yields more reliable results than frame-based schemes because the block decomposition process produces multiple samples. The research results indicate that increasing the number and size of blocks plays a crucial role in improving block-based steganalysis performance. However, if non-overlapping blocks are used, the relationship between these two parameters is inverted. This study proposed using overlapping blocks to resolve this contradiction, i.e., by increasing the overlapping level of the blocks, the number of blocks can be increased proportionally to the desired size.

The experimental results revealed that several parameters affected the performance of the proposed method and that there was a trade-off between these parameters and the complexity of the proposed method, making selecting these parameters a complex and challenging process. The outcomes showed that the concept of overlapping blocks improved the detection accuracy of techniques in the spatial and JPEG domains by more than 9%. In addition, one of the advantages of the proposed method is its comparable accuracy and lower computational complexity compared to state-of-the-art deep learning-based steganalysis.

Given the impact of the concept of overlapping blocks on the more precise discovery of steganographic methods and the growing popularity of CNN networks, these two concepts can be combined to create a more effective model. In this new model, rather than CNN using the entire image as input, overlapping blocks of the image that is more suitable based on the image texture can be selected and used as CNN input in a preprocessing step. Furthermore, we can expect the CNN network to succeed because the smaller input size reduces its complexity.

Author Contributions

All the authors participated in the conceptualization, implementation, and writing.

Acknowledgment

This work is completely self-supporting, thereby no any financial agency’s role is available.

Conflict of Interest

The author declares no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

DCT	Discrete Cosine Transform
JPEG	Joint Photographic Experts Group
DNN	deep neural networks
CNN	convolutional neural networks
ABC	Artificial Bee Colony
SVM	Support Vector Machine
TPs	True Positives
TNs	True Negatives
IQM	Image Quality Metrics
LSBR	LSB Replacement
LSBM	LSB Matching
TSVQ	tree-structured vector quantization
PQ	Perturbed Quantization
WOW	Wavelet Obtained Weights
S-UNIWARD	Spatial Universal Wavelet Relative Distortion
BPP	Bit Per Pixels
NOBS	Non-Overlapping Blocks-based Steganalysis
OBS	Overlapping Blocks-based Steganalysis
FS	Frame-based Steganalysis
SPAM	Subtractive Pixel Adjacency Matrix
SRM	Spatial Rich Model

References

[1] S. Dhawan , R. Gupta, "Analysis of various data security techniques of steganography: A survey," Inf. Secur. J., 30(2): 63–87, 2021.

[2] I. J. Kadhim, P. Premaratne, P. J. Vial, B. Halloran, "Comprehensive survey of image steganography: Techniques, evaluations, and trends in future research," Neurocomputing, 335: 299–326, 2019.

[3] V. Sabeti, M. Sobhani, S. M. H. Hasheminejad, "An adaptive image steganography method based on integer wavelet transform using genetic algorithm," Comput. Electr. Eng., 99: 107809, 202.

[4] R. Atta, M. Ghanbari, "A high payload data hiding scheme based on dual tree complex wavelet transform," Optik, 226: 165786, 2021.

[5] G. Paul, I. Davidson, I. Mukherjee, S. S. Ravi, "Keyless dynamic optimal multi-bit image steganography using energetic pixels," Multimed. Tools Appl., 76(5): 7445–7471, 2017.

[6] B. T. Hammad, I. T. Ahmed, N. Jamil, "A steganalysis classification algorithm based on distinctive texture features," Symmetry, 14(2): 236, 2022.

[7] M. Dalal, M. Juneja, "Steganography and steganalysis (in digital forensics): a cybersecurity guide," Multimed. Tools Appl., 80(4): 5723–5771, 2021.

[8] S. Cho, J. Wang, C. C. J. Kuo, B. H. Cha, "Block-based image steganalysis for a multi-classifier," in Proc. 2010 IEEE International Conference on Multimedia and Expo: 1457–1462, 2010.

[9] S. Cho, B. H. Cha, J. Wang, C. C. J. Kuo, "Performance study on block-based image steganalysis," in Proc. 2011 IEEE International Symposium of Circuits and Systems (ISCAS): 2649–2652.

[10] S. Cho, B. H. Cha, M. Gawecki, J. Kuo, "Block-based image steganalysis: Algorithm and performance evaluation," J. Vis. Commun. Image Represent., 24: 846–856, 2013.

[11] X. Lin, "Steganography and steganalysis," in Introductory Computer Forensics, Springer, Cham, 557–577, 2018.

[12] M. Broda, V. Hajduk, D. Levický, "Universal statistical steganalytic method," J. Electr. Eng., 68(2): 117–124, 2017.

[13] X. Y. Luo, D. S. Wang, P. Wang, F. L. Liu, "A review on blind detection for image steganography," Signal Processing, 88(9): 2138-2157, 2008.

[14] F. G. Mohammadi, H. Sajedi, "Region based Image Steganalysis using Artificial Bee Colony," J. Vis. Commun. Image Represent., 44: 214–226, 2017.

[15] A. Dehdar, A. Keshavarz, N. Parhizgar, "Image steganalysis using modified graph clustering based ant colony optimization and random forest," Multimed. Tools Appl., 1–18, 2022.

[16] J. Chen, W. Lu, Y. Fang, X. Liu, Y. Yeung, Y. Xue, "Binary image steganalysis based on local texture pattern," J. Vis. Commun. Image Represent., 55: 149–156, 2018.

[17] R. Wang, M. Xu, X. Ping, T. Zhang, "Steganalysis of JPEG images by block texture based segmentation," Multimed. Tools Appl., 74(15): 5725–5746, 2015.

[18] G. R. Suryawanshi, S. N. Mali, "Study of effect of DCT domain steganography techniques in spatial domain for JPEG Images steganalysis," Int. J. Comput. Appl., 127(6): 16–20, 2015.

[19] G. R. Suryawanshi, S. N. Mali, "Universal steganalysis using IQM and multiclass discriminator for digital images," Int. Conf. Signal Process. Commun. Power Embed. Syst.: 877–881, 2016.

[20] J. Zhu, X. Zhao, Q. Guan, "Detecting and Distinguishing Adaptive and Non-Adaptive Steganography by Image Segmentation," Int. J. Digit. Crime Forensics, 11(1): 62–77, 2019, doi: 10.4018/IJDCF.2019010105

[21] T. Pevny, J. Fridrich, "Merging markov and DCT features for multi-class JPEG steganalysis," Secur. Steganography, Watermarking Multimed. Contents IX, 6505: 650503, 2007.

[22] C. Chen, Y. Q. Shi, "JPEG image steganalysis utilizing both intrablock and interblock correlations," in Proc. IEEE Int. Symp. Circuits Syst.: 3029–3032, 2008.

[23] T. Pevný, P. Bas, J. Fridrich, "Steganalysis by subtractive pixel adjacency matrix," IEEE Trans. Inf. Forensics Secur., 5(2): 215-224, 2010.

[24] J. Kodovský, S. Binghamton, J. Fridrich, "Calibration revisited," in Proc. 11th ACM workshop on Multimedia and security: 63-74, 2009.

[25] J. Fridrich, J. Kodovsky, "Rich models for steganalysis of digital images," IEEE Trans. Inf. Forensics Secur., 7(3): 868–882, 2012.

[26] H. Zhang, F. Liu, Z. Song, X. Zhang, Y. Zhao, "AG-Net: An advanced general CNN model for steganalysis," IEEE Access, 10: 44116–44122, 2022.

[27] G. Xu, Y. Xu, S. Zhang, X. Xie, "SFRNet: Feature extraction-fusion steganalysis network based on squeeze-and-excitation block and RepVgg Block," Secur. Commun. Networks, 2021: 1-10, 2021.

[28] S. Wu, S. Zhong, Y. Liu, M. Liu, "CIS-Net: A novel CNN model for spatial image steganalysis via cover image suppression," arXiv preprint arXiv:1912.06540, 2019.

[29] B. Li, W. Wei, A. Ferreira, S. Tan, "ReST-Net: Diverse activation modules and parallel subnets-based CNN for spatial image steganalysis," IEEE Signal Process. Lett., 25(5): 650–654, 2018.

[30] S. Wu, S. H. Zhong, Y. Liu, "A Novel Convolutional Neural Network for Image Steganalysis with Shared Normalization," IEEE Trans. Multimed., 22(1): 256–270, 2017, doi: 10.48550/arxiv.1711.07306.

[31] J. Ye, J. Ni, Y. Yi, "Deep learning hierarchical representations for image steganalysis," IEEE Trans. Inf. Forensics Secur., 12(11): 2545–2557, 2017.

[32] G. Xu, H. Z. Wu, Y. Q. Shi, "Structural design of convolutional neural networks for steganalysis," IEEE Signal Process. Lett., 23(5): 708–712, 2016.

Biographies

Vajiheh Sabeti is an Assistant Professor of Engineering and Technology department at Alzahra University. She received her B.S. degree in Software Engineering in 2004 and her M.Sc. degree in Computer Architecture in 2007 and her Ph.D. degree in Computer Engineering in 2012 from the Electrical and Computer Engineering department of Isfahan University of Technology (IUT), Isfahan, Iran. Her research interests are softcomputing, image processing, and information hiding (steganography, watermarking).

- Email: v.sabeti@alzahra.ac.ir
- ORCID: 0000-0002-9985-9143
- Web of Science Researcher ID: AAD-1661-2022
- Scopus Author ID: 24470278600
- Homepage: https://staff.alzahra.ac.ir/vajiheh-sabeti/en/

How to cite this paper:
V. Sabeti, “An improved approach to blind image steganalysis using an overlapping blocks idea” J. Electr. Comput. Eng. Innovations, 11(2): 563-276, 2023.

DOI: 10.22061/jecei.2022.9241.592

URL: https://jecei.sru.ac.ir/article_1811.html