IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON RECENT ADVANCES IN VIDEO CODING AND SECURITY

Received January 7, 2019, accepted January 22, 2019, date of publication February 18, 2019, date of current version February 27, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2896781

Steganalysis of Content-Adaptive Steganography Based on Massive Datasets Pre-Classification and Feature Selection

JICANG LU[1], GANG ZHOU[1], CHUNFANG YANG[1], ZHENYU LI[2], AND MINGJING LAN[1]

[1]Zhengzhou Science and Technology Institute, Zhengzhou 450002, China
[2]Department of Computer Science, University of York, York YO 10 5GH, U.K.

Corresponding author: Chunfang Yang (chunfangyang@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61602508, Grant 61872448, Grant 61772549, and Grant U1736214.

ABSTRACT For current steganalysis of the image content-adaptive steganography, there are multiple problems to be improved, such as high difficulty and low accuracy, when detecting images with various contents and textures. For this problem, an improved steganalytic method is proposed in this paper based on the pre-classification and feature selection. First, using the features extracted based on the dependency analysis of image adjacent data, the images with various content and texture complexities are pre-classified as multiple clusters by the K-means algorithms. Then, the performance of existing various steganalytic features are analyzed for different clusters of images, and the optimal features for each cluster are selected for final classification. The experimental results show that the detection accuracy could be improved by the proposed method, and the rationality and availability are also verified. At the same time, the analysis and experimental results in this paper also show that the images with rich content and complex texture should be paid more attention both in steganography and in steganalysis.

INDEX TERMS Computer applications, computer security, data analysis, feature extraction, information security, multimedia communication, watermarking, steganalysis.

I. INTRODUCTION

Digital Steganography is a technique to embed confidential messages into redundant of multimedia files such as image, video, audio and text, and then transfer them through public channel for covert communication [1]. Generally, the redundancies are stochastic noise that, after embedding messages, statistical characteristics of the cover image would not be changed obviously [2]. Sometimes, the development of steganography could be considered as the decreasing of changing to actual redundant, then, the ability to defend steganalysis could be increased. Steganalysis, which is the inverse analysis technique of steganography, focuses on discovering the abnormal or artificial changing by analyzing the statistical characteristic of image data. The researching aspects of steganalysis include existence detection of confidential messages [3], embedding ratio estimation, extraction of messages [4], and etc. Therefore, developments of those two techniques are a game procedure in the past more than twenty years, and a great deal of outstanding theory or methods have been presented [5]–[8].

The appearing of image content-adaptive steganography based on the framework of uniting distortion function and STC (Syndrome-Trellis-Codes) [9], made the steganography technology step into a new stage. A lot of algorithms have been proposed based on this framework. For example, the typical algorithms used for spatial domain images include HUGO (Highly Undetectable steGO) [9], WOW (Wavelet Obtained Weights) [10], SUNIWARD (Spatial UNIversal WAvelet Relative Distortion) [11], MiPoD (Minimizing the Power of Optimal Detector) [12], CPP (Controversial Pixels Prior) [13], and etc. The typical algorithms used for JPEG images include J-UNIWARD (JPEG UNIversal WAvelet Relative Distortion) [11], SI-UNIWARD (Side-Informed UNIversal WAvelet Relative Distortion) [11], UED (Uniform Embedding Distortion) [14], and etc. These algorithms have made large difficult for steganalysis. In order to capture small changes caused by content-adaptive steganography, more steganalytic methods tend to extract rich model features, which is a popular direction at present [15].

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaoqing Pan.

2169-3536 © 2019 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

The typical rich model features include SRM (Spatial Rich Model) [16], SRMQ1 (SRM with Q1 strategy) [16], maxSRM (SRM with each bin of co-occurrences holds the sum of maximum values of the four embedding change probabilities at the corresponding residuals) [17], Adaptive SRM [18], PSRM (Projection SRM) [19], CC-JRM (Cartesian-calibrated JPEG Rich Model) [20], PHARM (PHase-Aware Projection Features) [21], DCTR (Discrete Cosine Transform Residuals) [22], GFR (Gabor Filter Residuals) [23], SCA (Selection Channel Aware) [24], and etc. At the same time, some steganalytic algorithms [25]–[27] were also designed based on advanced deep learning technique [28], 29], as well as steganographic algorithms [30]. Although the steganalytic algorithms based on these features perform well in some extent, there are still some problems [31], [32] such as low detection accuracy, poor robustness on detecting various types of images, and etc. Even when the deep learning technique is used, it is still hard to achieve ideal detection results.

This paper focuses on the detection of image content-adaptive steganography. Based on the analysis of dependency between the stego property and image content, a steganalytic framework based on massive datasets pre-classification and feature selection is presented for content-adaptive steganography. Using the current typical steganographic algorithms and steganalytic features, the efficiency and rationality of the proposed framework and detection algorithm are tested and verified by experiments. The obtained analysis results could not only help to improve the detection accuracy for content-adaptive steganography, but also could provide reference for improving security of steganographic algorithms in actual application.

The rest contents of this paper are arranged as follows: in section II, the limitation on improving the detection accuracy of content-adaptive steganography is presented by analyzing the stego property. In section III, the detection framework and algorithm procedure of the proposed method is introduced, which include the massive image datasets pre-classification based on complexity of image content, and the feature selection and assembling method. Section IV is about the experiments to test and verify the proposed method based on the popular image datasets, typical steganographic algorithms and steganalytic features. The paper is concluded in section V.

II. PROBLEMS STATEMENT

For the content-adaptive steganography, traditional steganalytic algorithms could hardly achieve effective detection. Although the methods based on rich model features are positive in some extent, it is still negative when detecting images with low embedding ratios or high complex texture. In order to seek the reasons of above problems, this paper will firstly analyze the principle of content-adaptive steganography and the relativity between steganography and image content in the following.

The main idea of the content-adaptive steganography based on uniting distortion function and STC includes two parts are the quantitative analysis of changing cost based on distortion function and the embedding based on STC. The distortion function aims to capture the changing of local or global characteristic after embedding by quantitative analysis, e.g., calculate the possible distortion of each elements after changing. According to the distortion, STC aims to determine the final elements to be changed by overall consideration, and then, the overall distortion could be minimized. Generally, the embedding change could be considered as additive noise, then, the noise of stego image consists of image original noise and additive stego noise. Therefore, the changing of cover image data could also be considered as the changing of cover image noise. For nature cover images, the noise intensity is usually associated with image content. For images with rich content and complex texture, the noise intensity is high. For images with simple content and texture, the noise intensity is low. Actually, the intensity of the stego noise is usually very low, and could be easily covered by original image noise with high intensity.

Then, it could also be analyzed from the point of view on adaptive steganalytic algorithms based on rich model features, which aims to capture the minor changes in steganography. At present, existing detection methods usually made no discrimination on images with various content and texture, and put them in the same dataset for training and classification. Then, the detection ability of a steganalytic algorithm is finite for images with various content and texture. If a group of features are sensitive to the stego noise in images with high intensity original noise, then, they could also be sensitive to that in images with low intensity original noise. However, if the features are sensitive to the stego noise in images with low intensity original noise, it is uncertainty that whether or not they are sensitive to the stego noise in images with high intensity original noise. Therefore, it is not an advisable decision to train and detect images by taking various intensity noises as a single datasets, and appropriate feature vector or components should be selected for each type of images [33], [34].

In addition, the following results could also be obtained according to above analysis. For a cover image I_1 with simple content and texture, the noise intensity is quantified as S_1. After embedding, the image became I_1', and the noise intensity became S_1' which satisfies the expression $S_1' > S_1$. At the same time, there is another cover image I_2 whose content and texture is slightly richer and more complex than I_1, then, there would be $S_2 > S_1$. Then, there is a probability that $S_1' \approx S_2$, and the rich model feature F_{R1}' extracted from I_1' will also be approximately equal to the rich model feature F_{R2} extracted from I_2. Finally, the stego image I_1' and cover image I_2 will be misjudged as the same class, which will decrease the detection accuracy.

In summary, it could be seen that, if the various content and texture of different images are not considered in steganalysis, e.g., training and classifying them in a single image dataset,

the detection accuracy will be decreased. This is one of the reasons that why the detection accuracy is poor, and is also the problem would be tried to solve in this paper.

III. THE PROPOSED FRAMEWORK AND DETECTION METHOD

In this section, the statistical feature that describes the image content complexity will be extracted based on the analysis of dependency between image adjacent data, and the feature are used for images pre-classification by K-means algorithm. Then, according to the pre-classification results, the phenomenon described in the last section is analyzed by experiments. At last, the detection framework based on pre-classification and feature selection is presented for steganalysis of content-adaptive steganography, and the detailed procedures are also described.

A. IMAGE PRE-CLASSIFICATION BASED ON CONTENT COMPLEXITY

Generally, there is strong dependency between adjacent data of digital image. However, the stochastic or irrelevant noise would be produced during the procedure of signal gathering, data quantizing and storing, which may destroy the strong dependency. In the area of computer vision, no matter for digital image or video, this type of characteristic is usually taken as a prerequisite on the further data analysis, such as noise analysis [35], motion estimation [36], [37], and etc. Therefore, the noise intensity of images could be reflected by analyzing the data dependency. This paper focuses on the analysis of digital images which is used in steganography. In this paper, the co-occurrence matrix (CM) of differences between image adjacent data is calculated to quantify the dependency of image data.

Denote the image with size $H \times W$ as I, and $I(h, w)$ is the value of pixel in location (h, w), $1 \le h \le H, 1 \le w \le W$. Denote the difference matrix along horizontal, vertical and diagonal directions as D^{hor}, D^{ver} and D^{diag}, respectively. The sizes of them are $H \times (W-1)$, $(H-1) \times W$ and $(H-1) \times (W-1)$, respectively. They could be calculated using following equations:

$$D^{hor} = I(1:H,\ 1:W-1) - I(1:H,\ 2:W) \quad (1)$$

$$D^{ver} = I(1:H-1,\ 1:W) - I(2:H,\ 1:W) \quad (2)$$

$$D^{diag} = I(1:H-1,\ 1:W-1) - I(2:H,\ 2:W) \quad (3)$$

On the basis of above differences matrixes, the CM for each direction could be obtained. It has been indicated by existing researches [35] that, CM of differences between neighboring pixels is center symmetric about (0, 0). Therefore, the CM of absolute value of the pixels differences could also reflect the distribution properties well. In addition, the statistical threshold is set to T in this paper. That is to say, when the absolute difference is larger than T, it will be cut off as T. Denote the CMs of the difference matrixes as M^{hor}, M^{ver} and M^{diag}, respectively. They could be calculated using following equations:

$$M^{hor}(a,b) = \frac{\sum_{i=1}^{H}\sum_{j=1}^{W-2}\delta(a, |D^{hor}(i,j)|)\delta(b, |D^{hor}(i,j+1)|)}{H \times (W-1)} \quad (4)$$

$$M^{ver}(a,b) = \frac{\sum_{i=1}^{H-2}\sum_{j=1}^{W}\delta(a, |D^{hor}(i,j)|)\delta(b, |D^{hor}(i+1,j)|)}{H \times (W-1)} \quad (5)$$

$$M^{diag}(a,b) = \frac{\sum_{i=1}^{H-2}\sum_{j=1}^{W-2}\delta(a, |D^{hor}(i,j)|)\delta(b, |D^{hor}(i+1,j+1)|)}{H \times (W-1)} \quad (6)$$

where $a, b = 1, 2, 3, \ldots, T$. $\delta(u, v) = 1$ if and only if $u = v$; otherwise, $\delta(u, v) = 0$.

According to above results, the final CM M used for pre-classification is calculated as follows:

$$M(a,b) = \frac{\left(M^{hor}(a,b) + M^{ver}(a,b) + M^{diag}(a,b)\right)}{3} \quad (7)$$

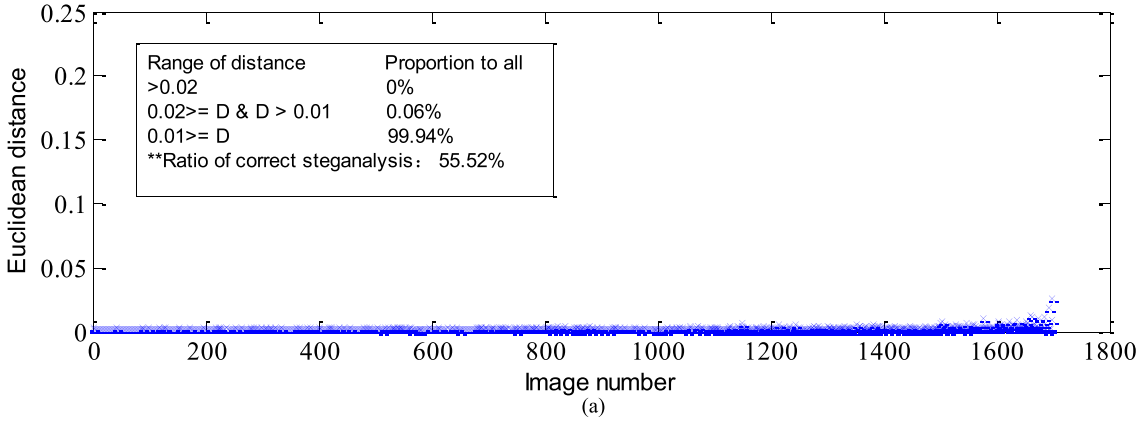
Taking M as the feature vector, the pre-classification for images with content of various complexities could be achieved using the classical K-means algorithm. On the basis of that, the detection accuracy could be improved by extracting appropriate features of each type of images for respective training and classification.

B. PROPERTY ANALYSIS OF CONTENT-ADAPTIVE STEGO IMAGES

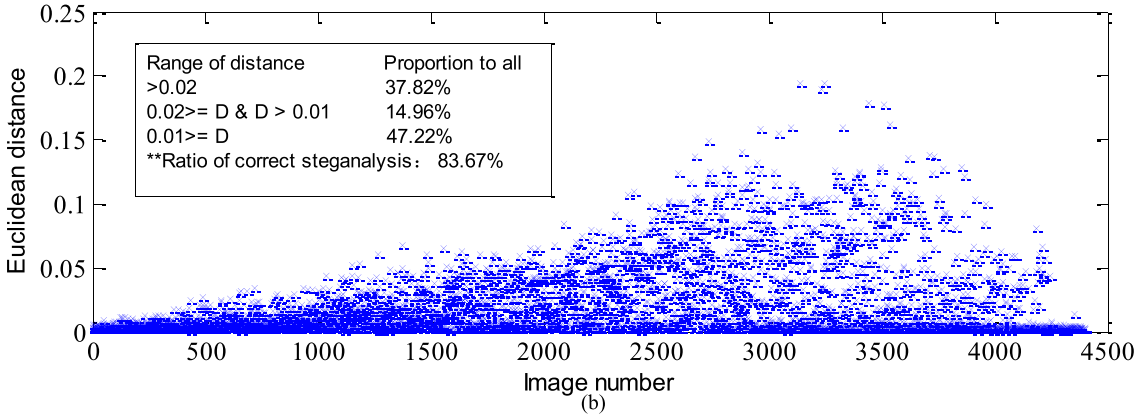
In this subsection, some experiments will be carried out. The images dataset is BOSSbase_1.01[1] with 10000 images and size 512×512, the typical steganographic algorithm is HUGO with embedding ratio 0.3 bpp (bits per pixel), and the rich model steganalytic feature is SRMQ1. Firstly, using the features CM with $T = 20$ and K-means algorithm to classify the cover images as 3 classes (clusters), the resulted quantity are 1699, 4405 and 3896, respectively. Then, denoting the SRMQ1 features of the n-th cover image and corresponding stego image in the i-th class as $F^C_{n,i}$ and $F^S_{n,i}$, respectively. The Euclidean distance $\Omega(F^C_{n,i}, F^S_{n,i})$ between them is calculated, $i = 1, 2, 3$. The results of each class of images are shown in Fig. 1(a), Fig. 1(b) and Fig. 1(c), respectively. Some typical sample images of each class are also shown.

It can be seen from the results in Fig. 1 that, the distance distribution between cover image features and stego image features of each class are different. The results in Fig. 1(a) show that, all the features distances of images with complex texture are very small, which indicates that this type of feature might be not sensitive to this class of images. Therefore, the ratio of correct steganalysis is only 55.52%, which is close to random guessing. The results in Fig. 1(b) and Fig. 1(c)

[1] BOSSbase: Available at: http://agents.fel.cvut.cz/stegodata/.2018.



(a)



(b)

FIGURE 1. The Euclidean distances and typical sample images of each cluster of images. (a). Results of the images with complex texture. (b). Results of the nature scene images with simple texture. (c). Results of the entity images with simple texture.

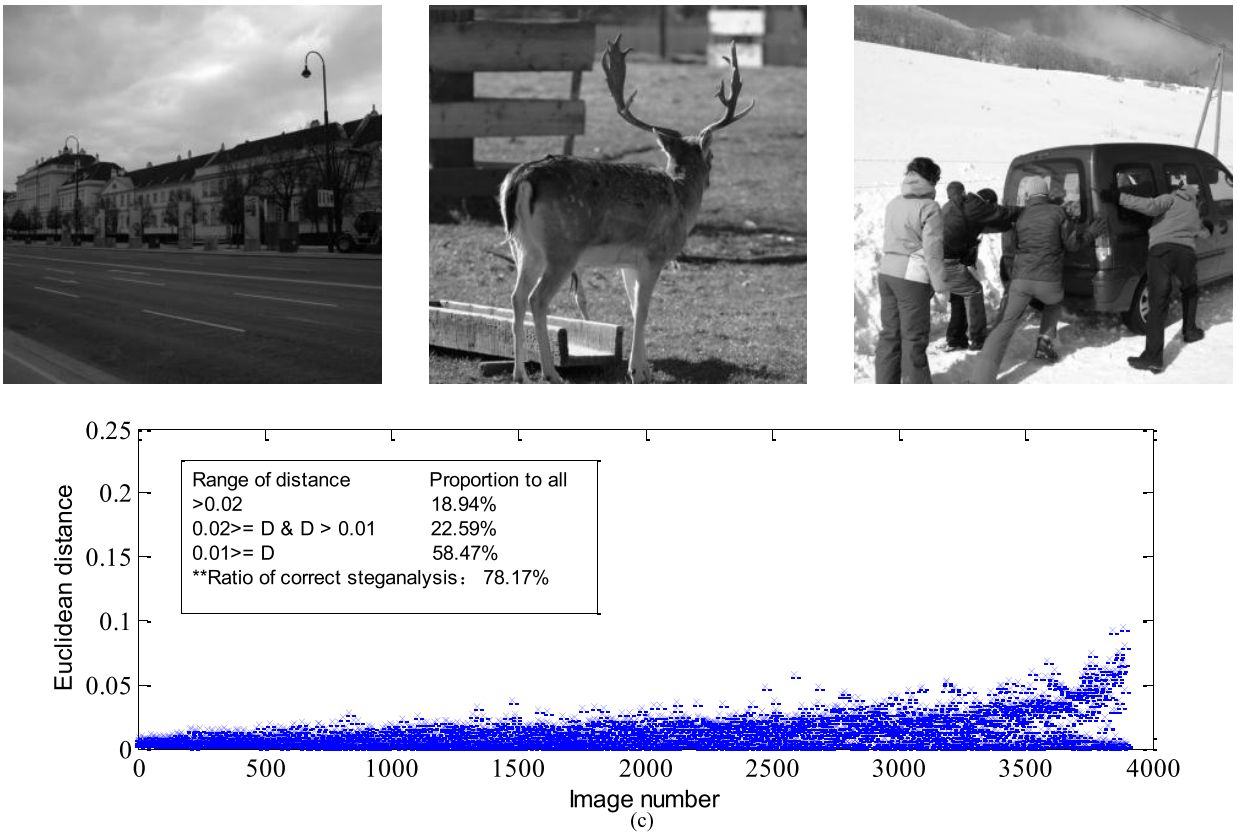


FIGURE 1. *(Continued.)* **The Euclidean distances and typical sample images of each cluster of images. (a). Results of the images with complex texture. (b). Results of the nature scene images with simple texture. (c). Results of the entity images with simple texture.**

show that, almost half of the features distances of these two classes of images are much larger, which indicates that this type of feature might be more sensitive to these two classes of images. Therefore, the ratios of correct steganalysis could reach 83.67% and 78.17%, respectively. It can be seen from part of sample images that, the images shown in Fig. 1(a) have richer content and more complex texture, which indicate that they might possess more intensity noise, and the minor stego noise might be covered. Then, the stego noise might not lead to obvious change of steganalytic feature, and distinguishability of the feature to cover and stego images is poor. In contrast, the images shown in Fig. 1(b) and Fig. 1(c) have simple content and texture, and there are also many smooth regions, which indicate that they might possess lower intensity noise. Then, the minor stego noise would be more obvious, the steganalytic feature would also be changed more seriously and have higher correct detection accuracy. The above results conform to the analysis in section II.

In addition, assembling separate detection results of the three classes of images, the global ratio of correct classification is 76.74%, which is higher than the detection results 75.53% before pre-classification. Although the improvement is not obvious, it could also be found that, most of the detection errors come from the images as shown in Fig. 1 (a). Then, if more appropriate and particular steganalytic features could be designed for those images, the overall detection accuracy could be improved further, which indicates that it would be feasible to improve the detection accuracy by the method based on pre-classification and feature selection proposed in this paper.

C. STEGANALYTIC FRAMEWORK AND METHOD BASED ON PRE-CLASSIFICATION AND FEATURE SELECTION

On the basis of above theoretical and experimental analysis, the proposed steganalytic framework based on pre-classification and feature selection for content-adaptive steganography is illustrated in Fig. 2. The framework consists of two stages, named training and detection. Each stage has two main steps. The training stage consists of pre-classification and feature selection, and the detection stage consists of pre-classification and detection. Then, the detailed applying procedure will also be described.

According to Fig. 2, the implementation procedures could be described as follows:

a. Training stage. The main purpose of this stage is to classify the training images with similar statistical property as the same cluster. Then, the property steganalytic features will be selected for each cluster of training images.

i. Pre-classification. When constructing training images set, the set always consists of various types of contents and textures. Then, they would be more similar to the condition in actual applications. For those images, extract

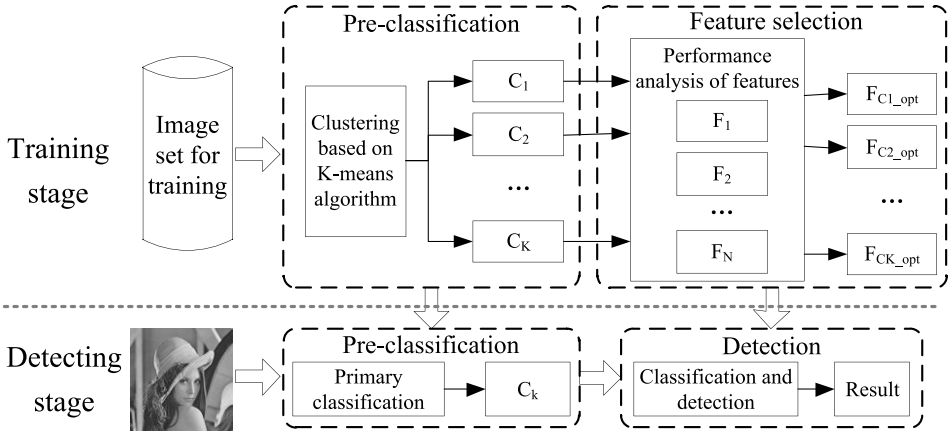


FIGURE 2. Framework of the proposed method based on pre-classification and feature selection.

pre-classification feature based on the relativity property which has been analyzed in section II. And then, the features are sent to K-means algorithm and the images are separated as K clusters $C_1, C_2, \ldots, C_k$.

ii. Feature selection. In order to describe the properties of the images better, multiple types of steganalytic features are usually compared and the feature with best performance will be selected for final steganalysis. For example, N types of typical feature extraction methods F_1, F_2, l, F_k are used to extract features of all the cover and corresponding stego images for the K clusters of images. Then, analyze the classification results of different features to each cluster of images, and select the feature with optimal performance for each cluster of images C_k, which is denoted as F_{Ck_opt}. At last, the selected feature will be used in actual steganalysis of the images in the C_k cluster.

b. Detection stage. The main purpose of this stage is to classify each testing image into the appropriate cluster trained in the training stage. And then, each image is detected using the appropriate feature trained and selected in the training stage.

i. Pre-classification. According to the K clusters of images obtained in training stage, analyze the similarity between the testing image and the training images, and pre-classify it as the most similar cluster C_k. All the testing images will be pre-classified, and each one belongs to only one cluster.

ii. Detection. According to the pre-classification result and, select the optimal steganalytic feature determined in training stage to classify the detected image, and output the final classification result. For example, if a testing image is pre-classified as the cluster of C_k, then, the feature F_{Ck_opt} of this image will be extracted. And then, the most possible correct classification result would be obtained.

IV. EXPERIMENTS AND RESULTS ANALYSIS

A. EXPERIMENTS SETUP

This subsection will introduce the image set, typical content-adaptive steganographic algorithms and rich model steganalytic features used in the experiments.

TABLE 1. Information about image datasets, steganographic and steganalytic algorithms.

Image Format	Steganographic algorithms	Embedding ratio	Rich model steganalytic features
JPEG with quality factor 75	nsf5 [38]	0.1 bpnzac	CC-JRM [20]
	J-UNIWARD [11]	0.3 bpnzac	DCTR [22]
	UED [14]		GFR [23]
	HUGO [9]	0.1 bpp	SRMQ1 [16]
pgm	S-UNIWARD [11]	0.3 bpp	maxSRMq2d2 [17]
	MiPod [12]		

The original cover images are from BOSSbase_1.01, which includes 10000 images with various content and texture. Format of the cover images is pgm with size 512×512, which could be directly used in the experiments of spatial domain steganography and steganalysis. For frequency domain images, the cover images are converted into JPEG image with quality factor 75 before used for steganography and steganalysis. For above two formats of images, the steganographic algorithms, embedding ratio and rich model steganalytic features are listed in Table 1, where bpnzac is the abbreviation of bits per no-zero AC coefficients.

B. EXPERIMENTAL RESULTS AND ANALYSIS

The detection framework and method proposed in section III will be tested based on the experiments setup in the last subsection. Firstly, the original cover images will be pre-classified. Then, analyze the performance of each steganalytic algorithm, and select the optimal feature for each cluster of images. At last, detect and present the overall classification results, and compare them with the detection results obtained by existing rich model steganalytic features. The classifier training strategy and method is same to that used in [16].

Firstly, the analysis and detection results after pre-classification for pgm images will be illustrated. During pre-classification, the threshold used in pre-classification feature extraction is set to 20, and the K in K-means algorithms are

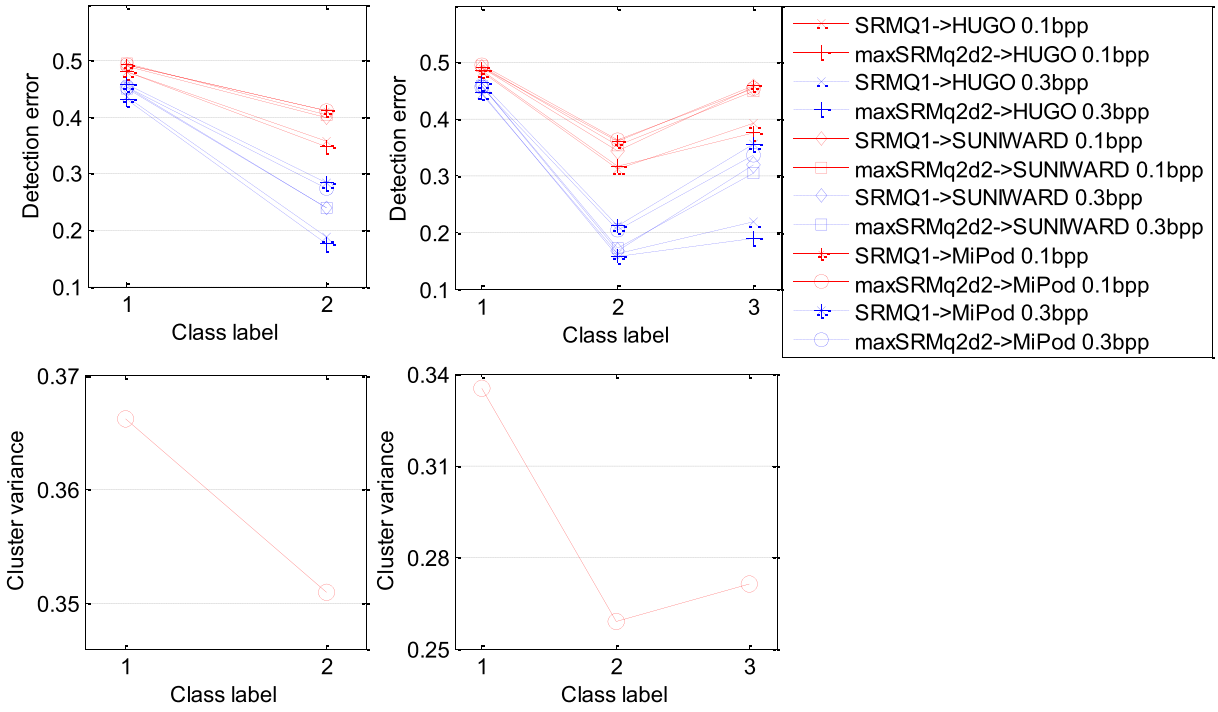


FIGURE 3. The congesting level and steganalytic results for each cluster of images after pre-classification.

set to 2 and 3, respectively. In fact, the parameter K is an empirical number, and the steganographer could set it according actual conditions. Generally, the more there are images for training, the bigger the number would be. For the clusters after pre-classification, analyze the congesting and scattering level based on the variance of pre-classification features. Generally, the smaller the variance is, the more congesting the images will be. Then, disturbance of the noise to cover image might be smaller, and the steganalytic performance might be better. In contrast, the steganalytic performance might be poor. For each cluster of images, detect them with each steganographic algorithm. The results are shown in Fig. 3. The possible relationship between the detection results and congesting level for each cluster will be analyzed later.

It could be seen from Fig. 3 that, the trend of steganalytic performance is similar to that of congesting level for all embedding ratios and steganographic algorithms, no matter the K in pre-classification is set to 2 or 3, or which steganalytic algorithm. It can be seen from the bottom left subfigure that, when the images are pre-classified as 2 clusters, congesting level of images with class label 1 is larger than that of images with class label 2, e.g., the images in cluster 1 are more scattering. The results in the upper left show that, all the detection error for cluster 1 are very high, which is little better than random guessing. In contrast, the detection errors for cluster 2 are lower, e.g., the correct detection accuracies are higher. The relationship reflected by the results shown in figures of the middle column is similar to that reflected by figures of the left column. Above results indicate that, the detection performance after pre-classification is associated with the congesting level, which could verify the previous analysis results.

On the basis of pre-classification results, the overall detection errors obtained by the proposed method for pgm and JPEG images are listed in Tables 2 and 3, respectively. For example, the N images are pre-classified as K clusters and their number of images are $n_1, n_2, \ldots n_K$, respectively, where $N = n_1 + n_2 + \ldots + n_K$. After detection, the errors of these clusters are $e_1, e_2, \ldots e_K$, respectively. Then, the overall detection error is $(n_1e_1 + n_2e_2 + \ldots + n_Ke_K)/N$.

At the same time, the results obtained by existing algorithms are also listed in the tables. The rows with F_orig at the first column denote the results without pre-classification. The rows with 2-Means at the first column denote the results after the images are classified as 2 clusters. The rows with 3-Means at the first column denote the results after the images are classified as 3 clusters. The rows with F_opt at the second column denote the detection results after selecting optimal features for each cluster of images.

It could be seen from the results in Tables 2 and 3 that, compared with the results of existing steganalytic features without pre-classification, all the detection errors of F_opt are lower, and almost all the detection errors of the existing feature after pre-classification are also lower, which indicate that the detection performance are improved by using the procedures of pre-classification and feature selection. Therefore, the rationality and availability of the proposed method could also be verified. In Table 2, for different types of images or different embedding ratios of stego images with the same steganographic algorithm, the superiorities of

TABLE 2. Comparison of detection errors before and after pre-classification for pgm images.

		HUGO		S-UNIWARD		MiPod	
		0.1	0.3	0.1	0.3	0.1	0.3
F_orig	SRMQ1	0.3900	0.2447	***0.4213***	0.2885	0.4302	0.3239
	maxSRMq2d2	***0.3843***	***0.2340***	0.4255	***0.2863***	***0.4298***	***0.3128***
2-Means	SRMQ1	0.3825	0.2386	0.4187	0.2843	0.4294	0.3205
	maxSRMq2d2	0.3753	0.2275	0.4222	0.2832	0.4277	0.3100
	F_opt	***0.3753***	***0.2275***	***0.4187***	***0.2832***	***0.4277***	***0.3100***
3-Means	SRMQ1	0.3732	0.2326	0.4119	0.2769	0.4215	0.3111
	maxSRMq2d2	0.3684	0.2200	0.4142	0.2724	0.4214	0.2979
	F_opt	***0.3664***	***0.2195***	***0.4090***	***0.2704***	***0.4192***	***0.2979***

TABLE 3. Comparison of detection errors before and after pre-classification for JPEG images.

		nsf5		J-UNIWARD		UED	
		0.1	0.3	0.1	0.3	0.1	0.3
F_orig	CCJRM	***0.2422***	***0.0181***	0.4728	0.3711	0.4335	0.2211
	DCTR	0.2823	0.0224	0.4491	0.2793	0.3863	0.1627
	GFR	0.3169	0.0489	***0.4262***	***0.2219***	***0.3534***	***0.1126***
2-Means	CCJRM	0.2395	0.0148	0.4739	0.3645	0.4322	0.2161
	DCTR	0.2826	0.0230	0.4467	0.2749	0.3844	0.1609
	GFR	0.3152	0.0499	0.4240	0.2176	0.3495	0.1107
	F_opt	***0.2395***	***0.0148***	***0.4240***	***0.2176***	***0.3495***	***0.1107***
3-Means	CCJRM	0.2403	0.0158	0.4729	0.3673	0.4308	0.2146
	DCTR	0.2813	0.0225	0.4481	0.2751	0.3842	0.1598
	GFR	0.3160	0.0490	0.4232	0.2184	0.3508	0.1100
	F_opt	***0.2403***	***0.0158***	***0.4232***	***0.2184***	***0.3508***	***0.1100***

different steganalytic feature are also different. For example, when detecting S-UNIWARD under different conditions, the feature SRMQ1 possesses the best performance for stego images with embedding ratio 0.1 bpp, but the feature maxSRMq2d2 possesses the best performance for stego images with embedding ratio 0.3 bpp. During feature selection, the proposed method could synthesize superiorities of all the two features, and then, could achieve the best performance under all the conditions, e.g., the detection errors are always the minimum among the three steganalytic algorithms. In Table 3, the performance of GFR is much better than other two features when detecting J-UNIWARD and UED, then, it is selected as optimal feature for all the clusters of images, which lead to that the detection accuracies of GFR and F_opt are equal to each other. Therefore, the detection errors are always the minimum among all the steganalytic algorithms. However, for traditional steganography nsf5, the feature CC-JRM possesses best performance, then, by using the methods proposed in this paper, it could be selected as optimal feature for nsf5. Therefore, no matter for traditional steganography or modern ones, the proposed method based on pre-classification and feature selection could always achieve the best detection performance.

The above results could verify the validity of analysis results in this paper. It is also verified rationality and availability to improve steganalytic performance of content-adaptive steganography by the proposed method based on pre-classification and feature selection. Besides, the analysis and experimental results in this paper indicate that, the images with rich content and complex texture should be paid more attention both in steganography and steganalysis. For steganography, the minor stego noise would be covered by high intensity noise of those images, and then would be harder to be captured. For steganalysis of content-adaptive steganography, it would be a main way to improve the detection accuracy by precisely capture the minor stego noise from images with high intensity noise.

V. CONCLUSIONS

For steganalysis of content-adaptive steganography, the detection performance might be poor when the content complexities of images are seriously different. To solve this problem, an improved steganalysis method based on pre-classification and feature selection is designed in this paper. Firstly, the content complexities of various images are analyzed, and then, the images are pre-classified using the feature of

co-occurrence matrix which could reflect the dependency of image adjacent data. Then, the performances of various steganalytic algorithms for different types of images are analyzed, and the features with best performance are selected for final classification, which could improve the overall steganalytic performance. At last, the proposed method is verified by experiments, and the suggestion on paying more attention to images with complex texture is also presented and analyzed.

REFERENCES

[1] J. Fridrich, *Steganography in Digital Media: Principles, Algorithms, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
[2] L. Xiang, Y. Li, W. Hao, P. Yang, and X. Shen, "Reversible natural language watermarking using synonym substitution and arithmetic coding," *Comput. Mater. Continua*, vol. 55, no. 3, pp. 541–559, 2018.
[3] A. D. Ker *et al.*, "Moving steganography and steganalysis from the laboratory into the real world," in *Proc. 1st ACM Workshop Inf. Hiding Multimedia Secur.*, 2013, pp. 45–58.
[4] C. Yang, X. Luo, J. Lu, and F. Liu, "Extracting hidden messages of MLSB steganography based on optimal stego subset," *Sci. China, Inf. Sci.*, vol. 61, no. 11, 2018, Art. no. 119103.
[5] A. Cheddad, J. Condell, K. Curran, and P. M. Kevitt, "Digital image steganography: Survey and analysis of current methods," *Signal Process.*, vol. 90, no. 3, pp. 727–752, Mar. 2010.
[6] X.-Y. Luo, D.-S. Wang, P. Wang, and F.-L. Liu, "A review on blind detection for image steganography," *Signal Process.*, vol. 88, no. 9, pp. 2138–2157, 2008.
[7] Y.-Q. Shi, X. Li, X. Zhang, H.-T. Wu, and B. Ma, "Reversible data hiding: Advances in the past two decades," *IEEE Access*, vol. 4, pp. 3210–3237, 2016.
[8] K. Karampidis, K. Kavallieratou, and G. Papadourakis, "A review of image steganalysis techniques for digital forensics," *J. Inf. Secur. Appl.*, vol. 40, pp. 217–235, Jun. 2018.
[9] T. Pevný, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *Proc. 12th Int. Workshop Inf. Hiding*, Calgary, AB, Canada, 2010, pp. 161–177.
[10] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *Proc. IEEE Int. Workshop Inf. Forensics Secur.*, Dec. 2012, pp. 234–239.
[11] V. Holub and J. Fridrich, "Digital image steganography using universal distortion," in *Proc. 1st ACM Workshop Inf. Hiding Multimedia Secur.*, 2013, pp. 59–68.
[12] V. Sedighi, R. Cogranne, and J. Fridrich, "Content-adaptive steganography by minimizing statistical detectability," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 2, pp. 221–234, Feb. 2016.
[13] W. Zhou, W. Zhang, and N. Yu, "A new rule for cost reassignment in adaptive steganography," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 11, pp. 2654–2667, Nov. 2017.
[14] L. Guo, J. Ni, and Y. Q. Shi, "Uniform embedding for efficient JPEG steganography," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 5, pp. 814–825, May 2014.
[15] Y. Yang, Y. Chen, Y. Chen, and B. Wei, "A novel universal steganalysis algorithm based on the IQM and the SRM," *Comput. Mater. Continua*, vol. 56, no. 3, pp. 261–272, 2018.
[16] J. Fridrich and J. Kodovský, "Rich models for steganalysis of digital images," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 868–882, Jun. 2012.
[17] T. Denemark, V. Sedighi, and V. Holub, "Selection-channel-aware rich model for steganalysis of digital images," in *Proc. IEEE Int. Workshop Inf. Forensics Secur.*, Dec. 2014, pp. 48–53.
[18] W. Tang, H. Li, W. Luo, and J. Huang, "Adaptive steganalysis based on embedding probabilities of pixels," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 4, pp. 734–745, Apr. 2016.
[19] V. Holub and J. Fridrich, "Random projections of residuals for digital image steganalysis," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 12, pp. 1996–2006, Dec. 2013.
[20] J. Kodovský and J. Fridrich, "Steganalysis of JPEG images using rich models," in *Proc. SPIE*, vol. 8303, p. 83030A, Feb. 2012.
[21] V. Holub and J. Fridrich, "Phase-aware projection model for steganalysis of JPEG images," *Proc. SPIE*, vol. 9409, p. 94090T, Mar. 2015.
[22] V. Holub and J. Fridrich, "Low-complexity features for JPEG steganalysis using undecimated DCT," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 2, pp. 219–228, Feb. 2015.
[23] X. Song, F. Liu, C. Yang, X. Luo, and Y. Zhang, "Steganalysis of adaptive JPEG steganography using 2D Gabor filters," in *Proc. 3rd ACM Workshop Inf. Hiding Multimedia Secur.*, 2015, pp. 15–23.
[24] T. D. Denemark, M. Boroumand, and J. Fridrich, "Steganalysis features for content-adaptive JPEG steganography," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 8, pp. 1736–1746, Aug. 2016.
[25] G. Xu, "Deep convolutional neural network to detect J-UNIWARD," in *Proc. 5th ACM Workshop Inf. Hiding Multimedia Secur.*, 2017, pp. 67–73.
[26] J. Zeng, S. Tan, B. Li, and J. Huang, "Large-scale JPEG image steganalysis using hybrid deep-learning framework," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 3, pp. 1200–1214, May 2018.
[27] M. Boroumand, M. Chen, and J. Fridrich, "Deep residual network for steganalysis of digital images," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1181–1193, May 2019. doi: 10.1109/TIFS.2018.2871749.
[28] D. Hu, L. Wang, W. Jiang, S. Zheng, and B. Li, "A novel image steganography method via deep convolutional generative adversarial networks," *IEEE Access*, vol. 6, pp. 38303–38314, 2018.
[29] R. Meng, S. G. Rice, and J. Wang, "A fusion steganographic algorithm based on faster R-CNN," *Comput. Mater. Continua*, vol. 55, no. 3, pp. 1–16, 2018.
[30] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[31] M. Boroumand and J. Fridrich, "Applications of explicit non-linear feature maps in steganalysis," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 4, pp. 823–833, Apr. 2018.
[32] L. Xiang, G. Zhao, Q. Li, W. Hao, and F. Li, "TUMK-ELM: A fast unsupervised heterogeneous data learning approach," *IEEE Access*, vol. 6, pp. 35305–35315, 2018.
[33] Y. Ma, X. Luo, X. Li, Z. Bao, and Y. Zhang, "Selection of rich model steganalysis features based on decision rough set α-positive region reduction," *IEEE Trans. Circuits Syst. Video Technol.*, to be published. doi: 10.1109/TCSVT.2018.2799243.
[34] J.-C. Lu, F.-L. Liu, and X.-Y. Luo, "Selection of image features for steganalysis based on the Fisher criterion," *Digit. Invest.*, vol. 11, no. 1, pp. 57–66, 2014.
[35] Q. Liu, A. Sung, and M. Qiao, "Neighboring joint density-based JPEG steganalysis," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 2, p. 16, 2011.
[36] Z. Pan, X. Yi, and L. Chen, "Motion and disparity vectors early determination for texture video in 3D-HEVC," *Multimedia Tools Appl.*, Nov. 2018, pp. 1–18. doi: 10.1007/s11042-018-6830-7.
[37] Z. Pan, J. Lei, Y. Zhang, and F. L. Wang, "Adaptive fractional-pixel motion estimation skipped algorithm for efficient HEVC motion estimation," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 14, no. 1, 2018, Art. no. 12.
[38] J. Fridrich, T. Pevný, and J. Kodovský, "Statistically undetectable JPEG steganography: Dead ends challenges, and opportunities," in *Proc. 9th ACM Workshop Multimedia Secur.*, 2007, pp. 3–14.

JICANG LU was born in Hebei, China, in 1985. He received the B.S., M.S., and Ph.D. degrees in computer science and technology from the Zhengzhou Science and Technology Institute, Zhengzhou, China, in 2007, 2010, and 2014, respectively, where he has been an Assistant Professor, since 2014.

His current research interests include digital image steganography and steganalysis, intelligent data analysis, and knowledge graph embedding and inference.

GANG ZHOU was born in Henan, China, in 1974. He received the B.S. and M.S. degrees in computer science and technology from the Zhengzhou Science and Technology Institute, Zhengzhou, China, in 1996 and 1999, respectively, and the Ph.D. degree in computer science and technology from Beihang University, Beijing, China, in 2006.

From 2007 to 2017, he was an Associate Professor with the Zhengzhou Science and Technology Institute, where he has been a Professor, since 2018. His research interests include intelligent data analysis, social networks, and big data mining.

CHUNFANG YANG was born in Fujian, China, in 1983. He received the B.S., M.S., and Ph.D. degrees in computer science and technology from the Zhengzhou Science and Technology Institute, Zhengzhou, China, in 2005, 2008, and 2012, respectively, where he was an Assistant Professor, from 2012 to 2016, and has been an Associate Professor, since 2017.

His research interests include information security, digital steganography, and steganalysis.

ZHENYU LI was born in Luoyang, China, in 1989. He received the B.S. degree in information computing sciences from the Hefei University of Technology, Hefei, China, in 2011, the M.S. degree in computer application technology from the Zhengzhou Institute of Information Science and Technology, Zhengzhou, China, in 2014, and the Ph.D. degree in computer science from the University of York, York, U.K., in 2018. His research interests include 3-D Information hiding, steganalysis, and machine learning.

MINGJING LAN was born in Anhui, China, in 1982. He received the B.S., M.S., and Ph.D. degrees in computer science and technology from the Zhengzhou Science and Technology Institute, Zhengzhou, China, in 2003, 2006, and 2013, respectively, where he was an Assistant Professor, from 2009 to 2015, and has been an Associate Professor, since 2016.

His research interests include big data analysis and knowledge graph.