IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 20, 2019, accepted November 17, 2019, date of publication November 22, 2019,
date of current version December 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2955113

Towards Improved Steganalysis: When Cover Selection is Used in Steganography

ZICHI WANG[1], SHUJUN LI[2,3], (Senior Member, IEEE), AND XINPENG ZHANG[1], (Member, IEEE)

[1]Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Joint International Research Laboratory of Specialty Fiber Optics and Advanced Communication, Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai 200444, China
[2]School of Computing, University of Kent, Canterbury CT2 7NF, U.K.
[3]Kent Interdisciplinary Research Centre in Cyber Security (KirCCS), University of Kent, Canterbury CT2 7NF, U.K.

Corresponding authors: Zichi Wang (wangzichi@shu.edu.cn) and Shujun Li (S.J.Li@kent.ac.uk)

This work was supported in part by the Natural Science Foundation of China under Grant U1636206 and Grant 61525203.

ABSTRACT This paper proposes an improved steganalytic method when cover selection is used in steganography. We observed that the covers selected by existing cover selection methods normally have different characteristics from normal ones, and propose a steganalytic method to capture such differences. As a result, the detection accuracy of steganalysis is increased. In our method, we consider a number of images collected from one or more target (suspected but not known) users, and use an unsupervised learning algorithm such as k-means to adapt the performance of a pre-trained classifier towards the cover selection operation of the target user(s). The adaptation is done via pseudo-labels from the suspected images themselves, thus allowing the re-trained classifier more aligned with the cover selection operation of the target user(s). We give experimental results to show that our method can indeed help increase the detection accuracy, especially when the percentage of stego images is between 0.3 and 0.7.

INDEX TERMS Cover selection, steganography, steganalysis, clustering.

I. INTRODUCTION

Steganography is the art of covert communication, aiming to transmit data secretly through public channels without drawing suspicion, and steganalysis aims to disclose the secret transmission by analyzing suspected media [1].

Modern steganalytic methods use supervised machine learning to investigate the models of the covers and the stegos. Features are extracted from a set of images to train a common steganalytic classifier, which is then used to distinguish real stego images from normal (cover) images without any hidden information [2], [3]. The ensemble classifier [4] is widely used to enhance the performance by using multiple classifiers. The feature extraction and machine learning based steganalysis has been proved to be efficient.

The most popular feature set is SRM (Spatial Rich Model) [5], which are the fourth order co-occurrence matrices for describing the dependencies among different pixels. After SRM, some improved feature extraction methods are proposed [6], [7]. In PSRM (Projections of Spatial Rich Model) [6], neighboring residual samples are projected onto a set of random vectors and the histograms of the projections are taken as the feature. The feature set maxSRMd2 [7] is a variant of SRM that makes use of the modification probabilities of cover elements during data embedding, which is called probabilistic selection channel. Recently, deep learning based steganalysis has also achieved good performances with enough training data [8]–[10].

To resist steganalysis, in modern steganography, the additive distortion between the cover and the stego needs minimizing to leave minimum statistical traces. One way of doing this is to use the syndrome trellis coding (STC) [11] with a user-defined distortion function, e.g., SUNIWARD [12], WOW [13], HILL [14] for spatial images, and JUNIWARD [12], UED [15], UERD [16], HDS [17] for JPEG images. In addition, the security performance can be significantly improved by the selection of the cover when there are a number of candidate images available to the user, as shown in Fig. 1. In this case, the most suitable images can be selected to minimize the detectability of the hidden information. Researchers have proposed many methods for cover

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

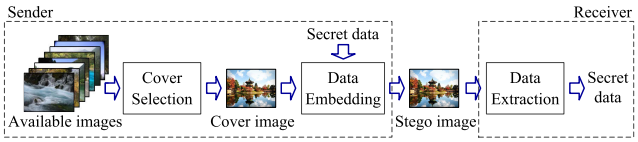

FIGURE 1. Cover selection based steganography.

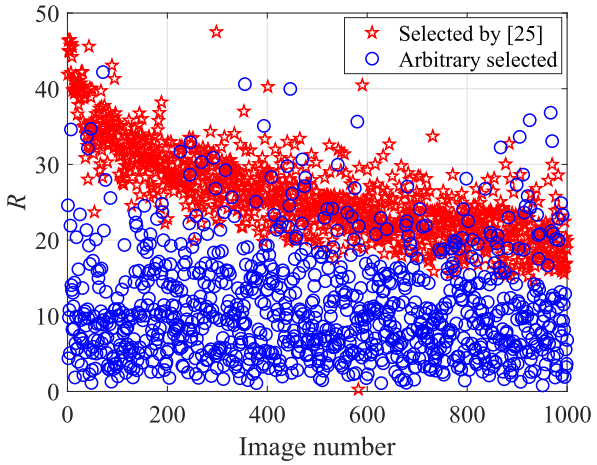

FIGURE 2. Residual values of the [25] selected and arbitrarily selected images.

selection [18]–[25]. The cover selection step is normally not considered in modern steganalysis methods, which explains why cover selection methods can help reduce detectability. However, if we consider the user behavior in cover selection into the steganalysis process, it may be possible to exploit some cover selection methods for improving performance of steganalysis.

Our proposed steganalytic method is based on the observation that covers selected for hiding information normally have different characteristics from normal ones (shown in Fig. 2), and such differences can be captured by a simple unsupervised clustering algorithm such as k-means. We assume that the image source of the user of steganographer is different with that of the steganalyst. This assumption is reasonable since the available images of steganographer and steganalyst are different. Some steganalytic methods focus on decreasing the drop in detection accuracy caused by image source mismatch [26]–[29], but these methods do not consider the cover selection operation for steganalysis. Based on this assumption, we can use a clustering algorithm to help re-train a machine learning based classifier to adapt to the cover selection operation of the target suspected user(s). The re-training is done via pseudo-labels derived from the predicted labels of the pre-trained classifier and those from the clustering algorithm, which are used to get new pseudo-training samples for improving the detectability. Our experimental results showed an improved detectability of the proposed steganalytic method. The reason why this worked is because the re-training process based on pseudo-labels can help reduce the mismatch of image sources between the user of steganography and the steganalyst, which is essentially an adaptation to the cover selection operation of the target user(s).

The rest of the paper is organized as follows. The next section introduces some related work. Our proposed method is explained in detail in Section III. Experimental results of the proposed method are given in Section IV. The last section concludes the paper.

II. RELATED WORK

In this section, some related work about the use of cover selection in steganography is introduced firstly. Then the differences between our method and some similar steganalytic methods are clarified.

Most of current cover selection methods select cover images empirically [18]–[22]. In [18], a cover selection method is proposed for JPEG steganography. Firstly, the coefficients which will be utilized by the embedding process (called changeable DCT coefficients) are counted. Then the images with a larger number of changeable DCT coefficients and higher quality factor are selected as covers for secret data carrying. In [19], the JPEG images with high visual quality (such as high quality factor) are regarded as suitable covers for embedding. The authors suggested that the images with poor quality not only negatively affect the capacity but also provoke special attention.

For uncompressed image, suitable images are selected according to image texture and complexity [20]–[22], since images with higher complexity have more details and human vision system is less capable to detect minor modifications. In [20], the secret information is also an image. The blocks of secret image are compared with the blocks of cover images, and then the images with most similar blocks to those of the secret image are selected as covers. Where the similarity of blocks is evaluated by the mean, variance, skewness of 2×2 sub-blocks, and the neighborhood information. In [21], image complexity is measured by visual quality (such as PSNR) and amount of changes on a stego-image since smaller amount of changes means higher visual quality and lower steganalysis detectability. Then the images with more complex texture are preferentially selected. The authors of [22] proposed to use spatial information which calculated from image residuals to measure image complexity. The spatial information based image complexity is modelled by fuzzy logic which finds the images that yield least detectable stego image.

In [23], the relationship between image characteristics (variance, complexity, entropy, histogram and function of histogram) and steganography performance (relative entropy and change of histogram) are explored, and then the images with mild histogram are regarded as suitable covers. A unified measure to evaluate the hiding ability of a cover image is proposed in [24] based on Fisher Information Matrix and Gaussian Mixture Model. Cover images are represented by the Gaussian mixture model firstly. Then the Fisher Information Matrix of cover image is calculated and mapped into a real value to evaluate the hiding ability. But the employed model is not able to describe natural image precisely. In [25], the first-order derivative of steganographic distortion of a

single cover is proved monotonically increasing with the value of payload increasing. In addition, it is proved that first-order derivative of steganographic distortion of covers that selected from a given set should be equal. Based on the deductions, the images with minimal total steganographic distortion are selected as covers, which results in high undetectability of steganography.

As a general approach, cover selection has a theoretical flaw that selected cover images (as a subset of all possible images) will always have some statistical properties different from the whole set of all possible images [30], [31]. The flaw will be shown in Section III-A. This would lead to the development of steganalysis methods that can break any cover selection methods. However, to the best of our knowledge, there is no publicly reported work about steganalysis against cover selection based steganography. Therefore, it remains unknown how the theoretical flaw can lead to practical attacks. The above-mentioned flaw also exists in some coverless and cover-generation based steganographic methods [32]–[34]. In these methods, the cover image is generated instead of using existing images. This generation should also in principle lead to different statistical properties from other existing images. Therefore, it is possible that our method can be extended to attack coverless and cover-generation based steganographic methods.

As a similar kind of steganalysis, pooled steganalysis [35]–[39] aims to group a set of clues in order to detect the use of embedding. Some pooled steganalytic methods [35], [36], [38] assumed that the steganalyst is monitoring a number of users, with multiple innocent users and some potentially genuine users. To determine who are the genuine users, it is assume that their behaviors significantly deviate from the majority of innocent users. Based on this assumption, the genuine users can be recognized by unsupervised clustering algorithms. Other pooled steganalytic methods [37], [39] consider the scenario with only one user and with the use of a single image dectector, also aims to find the users of steganography.

Different from these pooled steganalytic methods which work at the user level, our proposed behavioral steganalysis focus on the image level, which aims to find the stego-images among a number of clear images. With no doubt, the detection results on images made by our method can also contribute to the detection at the user level.

III. PROPOSED METHOD

In this section, we demonstrate the distribution separation of existing cover-selection methods firstly, and then propose our steganalytic method to capture this separation.

A. DISTRIBUTION SEPARATION OF COVER SELECTION

As mentioned in Section II, the statistical properties of selected cover images are different from other images that are not selected. To demonstrate this distribution separation, we conducted a group of experiments over the image dataset BOSSbase ver. 1.01. We selected 1,000 images from BOSSbase ver. 1.01 using the method in [25], then calculated the average value R of horizontal and vertical residual of each image, defined by Eq. (1), where $x(i,j)$ is the (i,j)-th pixel value of an image sized $M \times N$. For comparison, another set of 1,000 images were randomly selected from BOSSbase ver. 1.01 as the set of all possible images.

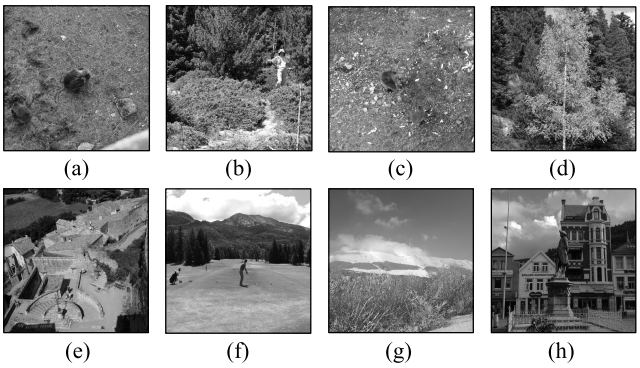

FIGURE 3. Demonstration of images (a)∼(d) selected by the method in [25], and images (e)∼(h) selected arbitrarily.

$$R = \frac{1}{MN}\left[\sum_{i=1}^{M}\sum_{j=1}^{N-1}|x(i,j)-x(i,j+1)| + \sum_{i=1}^{M-1}\sum_{j=1}^{N}|x(i,j)-x(i+1,j)|\right] \quad (1)$$

The comparison of residual values between the selected cover images with the arbitrary images is shown in Fig. 2. It can be seen that the residual distribution of the images selected by [25] is clearly different from the distribution of the set of all possible images. As a general tendency, the images selected by [25] contain larger residual values (i.e., more complex texture), as shown in Fig. 3, therefore the modification traces made by steganography can be concealed effectively. While the larger residual values are beneficial for embedding, the different statistical properties of the selected cover images can also lead to new risks of steganalysis. This flaw also exists in other cover selection methods, which has been verified in some other related work [30], [31]. Therefore, this statistical abnormality is a universal phenomenon of cover selection methods. As a result, it can be employed by steganalysis, which will be demonstrated in this paper. To measure the statistical abnormality, other metrics can be employed, e.g., deflection coefficient [40]. In this paper, we do not calculate the statistical abnormality accurately, since the statistical abnormality can be observed clearly in Fig. 2.

B. PROPOSED STEGANALYTIC METHOD

In this paper, we provide experimental evidence of a practical and effective steganalysis on a specific cover selection method, but the idea could be extended to break other cover selection methods without any change. Our proposed

steganalysis works by capturing the expected statistical differences by an unsupervised clustering algorithm such as k-means, given a set of images from one or more suspected users of a cover selection based steganographic system. The clustering algorithm helps re-train a machine learning based steganalytic classifier so that the latter can adapt to the cover selection behavior of the target user(s). The actual effect of the re-training is to reduce the mismatch between the image sources of the target user(s) and those of the steganalyst, therefore increasing the detection accuracy of the re-trained classifier.

Before explaining the proposed method in greater detail, let us clarify the scenarios. First, the steganalyst uses a set of his own images (e.g., a standard test image database) as the training set to obtain a base-line classifier. Subsequently, the steganalyst collects a number of images from one or more users, some of whom are suspected to be using steganography to hide secret information in some of the collected images. The task of the steganalyst is to detect which images in the collected set are stegos. We assume some cover selection method is used by the user(s) of steganography to make steganalysis harder.

The images collected are marked "clear" or "stego" by the base-line classifier and marked "class 1" or "class 2" by a clustering algorithm (we use the popular k-means algorithm as an example, other alternatives can also work). Based on the assumption that the clustering algorithm can capture some characteristics of the selected covers for steganography, we hypothesize that either class 1 or 2 contains more stego images. Based on this assumption and another one that the base-line classifier performs relatively well (which is generally true for a good base-line classifier), we can combine the two sets of labels to create new pseudo-labels for some images, which can be used to re-train the base-line classifier so that it can hopefully adapt itself more towards the genuine cover selection behavior of the target user(s). The re-trained classifier is used to re-classify all the images collected to produce the final prediction labels. The re-training is done separately for different sets of target images from the base-line classifier. More details are given below.

Assume that n images are collected for detection, which include k stego-images. Denote the n images by $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$. After classified by the base-line classifier, some images are marked as "clear", denoted by $\mathbf{X}_i^c$, and the others are marked as "stego", denoted by $\mathbf{X}_i^s$, $i \in \{1, 2, \ldots, n\}$. Similarly, some images are marked as "class 1" by the clustering algorithm, denoted by $\mathbf{X}_i^{(1)}$, and the other as "class 2", denoted by $\mathbf{X}_i^{(2)}$. As shown in Fig. 4, each image is marked with two labels. Our task now is to refine the classification labels from the base-line classifier based on the clustering based labels.

Denote images with labels "clear" and "class 1", "clear" and "class 2", "stego" and "class 1", 'stego" and "class 2" as $\mathbf{X}_i^{c(1)}$, $\mathbf{X}_i^{c(2)}$, $\mathbf{X}_i^{s(1)}$, $\mathbf{X}_i^{s(2)}$, respectively. Denote the corresponding numbers of images with the four different sets of

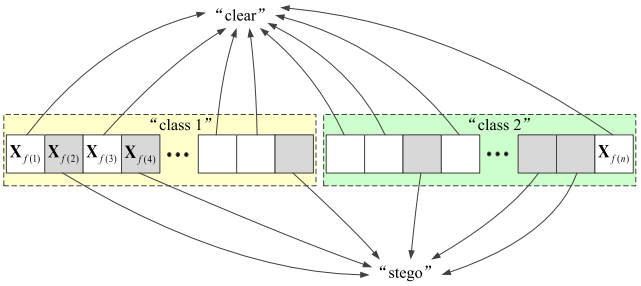

FIGURE 4. Labels of images for detection.

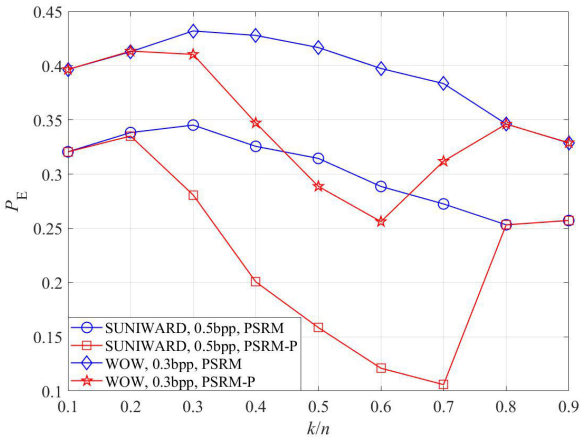

FIGURE 5. Detectability w.r.t. the value of k/n when $n = 10,000$.

labels by $\gamma^{c(1)}$, $\gamma^{c(2)}$, $\gamma^{s(1)}$, $\gamma^{s(2)}$, respectively. It is clear that $\gamma^{c(1)} + \gamma^{c(2)} + \gamma^{s(1)} + \gamma^{s(2)} = n$.

Based on these labels, we propose an image selection strategy to identify images with potentially more reliable pseudo-labels "clear" or "stego" for re-training the base-line classifier towards the images collected, so that the re-trained classifier can hopefully capture the specific characteristics of the images under inspection, i.e., the cover selection behavior of the target user(s). Note that the new labels are pseudo-labels (i.e., no human efforts are involved), so applying the re-trained classifier to those images used in the re-training process does not involve circular reasoning.

The main idea behind the image selection strategy is the following: when the base-line classifier and the clustering algorithm make the same judgment, the labels they agree are likely more reliable and can capture the characteristics of the stego and normal images for the specific collection better than the standard training database the steganalyst originally used to train the base-line classifier. Since the clustering algorithm does not produce "stego" and "clear" labels, we consider the majority labels in "class 1" and "class 2" as agreed labels if the two clustering-based classes contain different majority labels (i.e., one contains more "stego" labels and the other contains more "clear" labels, or vice versa). The above process can be formulated based on the following three inequalities, where Inequality (3) checks the two clustering based classes do contain different majority labels, and the

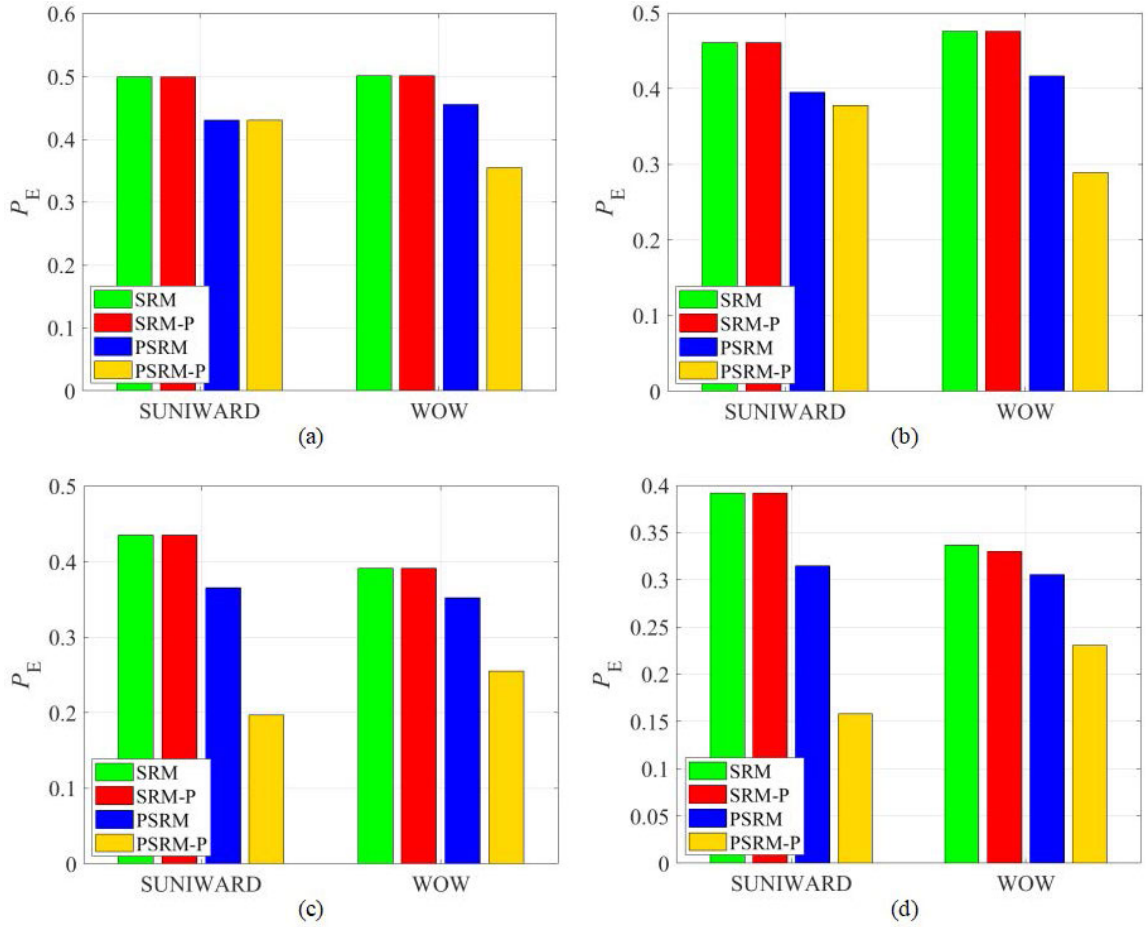

FIGURE 6. Detection error comparisons between SRM, PSRM and the improved versions using the proposed steganalytic method for $n = 10{,}000$ and $k = 5{,}000$ using SUNIWARD and WOW with (a) 0.2 bpp; (b) 0.3 bpp; (c) 0.4 bpp; (d) 0.5 bpp.

other two equations are the conditions corresponding to the two different mappings from the clustering labels to the labels of the base-line classifier. In detail, $\mathbf{X}_i^{c(1)}$ and $\mathbf{X}_i^{s(2)}$ will be added into the training samples if Inequality (4) is satisfied. That means, there are more "clear" labels than "stego" labels in "class 1", so that the images with labels "clear" in "class 1" and the images with labels "stego" in "class 2" should be added into the training samples. On the contrary, $\mathbf{X}_i^{c(2)}$ and $\mathbf{X}_i^{s(1)}$ will be added into the training samples if Inequality (4) is satisfied.

$$(\gamma^{c(1)} - \gamma^{s(1)})(\gamma^{c(2)} - \gamma^{s(2)}) < 0 \tag{2}$$

$$\gamma^{c(1)} > \gamma^{s(1)} \tag{3}$$

$$\gamma^{c(1)} < \gamma^{s(1)} \tag{4}$$

The proposed image selection strategy for re-training the base-line classifier can be summarized as Algorithm 1.

Algorithm 1 The Algorithm for Selecting Images With New Pseudo-Labels.

```
Mark the images collected with labels 'clear'' or ''stego'' using the base-line classifier;
Mark the images collected with labels ''class 1'' or ''class 2'' using a clustering algorithm;
if Inequality (3)is satisfied then
    if Inequality (4) is satisfied then
        Add X_i^{c(1)} and X_i^{s(2)} into the training samples;
        break;
    end if
    if Inequality (4) is satisfied then
        Add X_i^{c(2)} and X_i^{s(1)} into the training samples;
    end if
end if
```

IV. EXPERIMENTAL RESULTS

A. EXPERIMENT SETUP

The image datasets employed in our experiments are BOSSbase ver. 1.01 [41] that contains 10,000 uncompressed grayscale images sized 512×512, and UCID [42] that contains 1,338 uncompressed color images sized 512×384. The 1,338 images in UCID were transformed into grayscale images, and then used to train the base-line classifier. The 10,000 images in BOSSbase ver. 1.01 were used to sample different sets of images collected from suspected user(s) ($n \leq 10,000$). In detail, n images are arbitrarily selected from BOSSbase ver. 1.01 as the collected images for detection, and then k images are selected from the obtained n images by

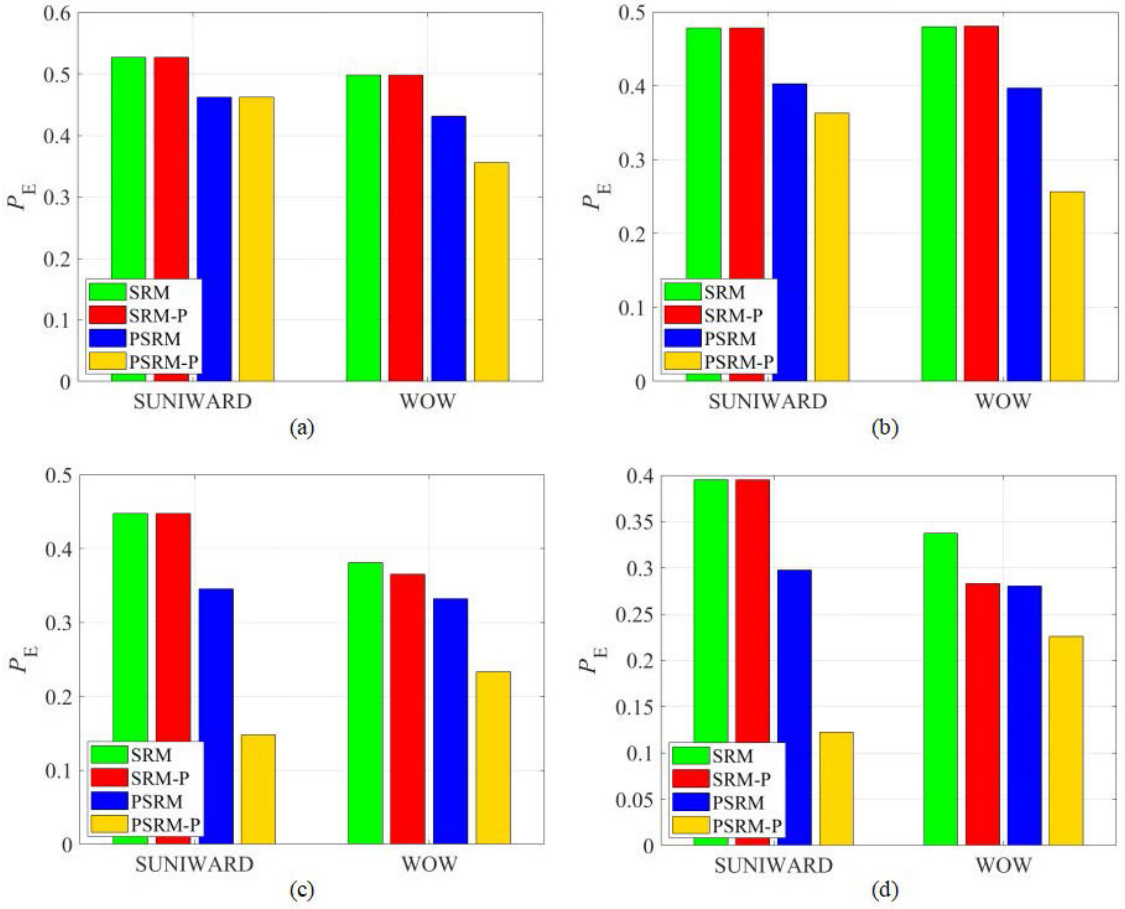

FIGURE 7. Detection error comparisons between SRM, PSRM and the improved versions using the proposed steganalytic method for $n = 10{,}000$ and $k = 6{,}000$ using SUNIWARD and WOW with (a) 0.2 bpp; (b) 0.3 bpp; (c) 0.4 bpp; (d) 0.5 bpp.

one of the existing cover selection methods for stego-images production.

For steganography, the popular methods SUNIWARD [12] and WOW [13] were used for embedding (to produce the stego images of the steganographer). For cover selection, the method in [25] was used to select the most suitable k covers since it achieves high undetectability against modern steganalysis.

For steganalysis, the popular feature extraction methods SRM [5] and PSRM [6] are used as benchmark. We employ the ensemble classifier proposed in [4] to measure the property of feature sets. For the clustering algorithm, we use k-means [43], one of the most used clustering methods.

The criterion of evaluating the performance of feature sets is the minimal total error P_E with identical priors achieved on the testing sets [4], defined by

$$P_E = \min_{P_{FA}} \left(\frac{P_{FA} + P_{MD}}{2} \right), \tag{5}$$

where P_{FA} is the false alarm rate and P_{MD} the missed detection rate. The performance is evaluated using the average of P_E over ten random tests. For each test, a random set of n images were selected out of the 10,000 BOSSbase database and k images out of the n ones were used as the covers. Different values of n and k were tested to check the performance under different scenarios.

B. DETECTABILITY

Since not all n images are stego images, the detectability of steganalysis depends on the ratio of stego images (k/n). The relationship between detectability and the value of k/n is shown in Fig. 5, where stego images were obtained using SUNIWARD with 0.5 bpp (bit per pixel) and WOW with 0.3 bpp, respectively. "PSRM-P" in the legend refers to the proposed method with the re-trained classifier, while "PSRM" refers to the base-line classifier.

It can be seen that the advantage of the proposed method becomes obvious when the value of k/n is between 0.3 and 0.7, and reaches the maximum around 0.6 and 0.7. On the one hand, it is reasonable that a larger ratio of stego images results in higher detectability, since the detectable trace made by steganography is in proportion to the quantity of embedded data. On the other hand, a larger ratio of stego images results in more unbalance quantity between clear and stego images when this ratio is larger than 0.5. This unbalance is unfavourable to the proposed method which employs clustering. Therefore, the optimal value of k/n for the proposed method is larger than 0.5 but not close to 1.

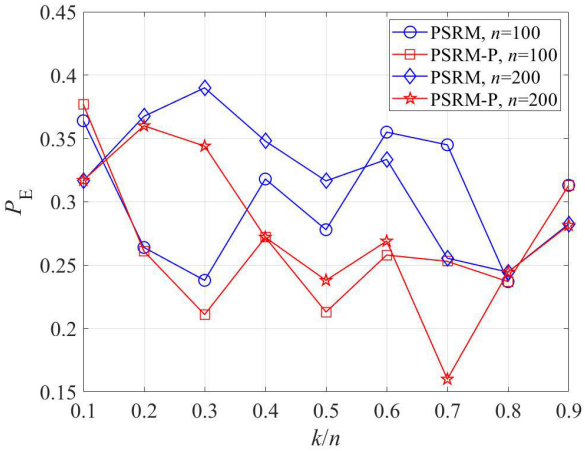

FIGURE 8. Detection errors for smaller image sets ($n = 100$ and $n = 200$).

The comparisons of detection error P_E with feature sets SRM, PSRM and the improved versions "SRM-P" and "PSRM-P" using the proposed method with $k = 5{,}000$ and $k = 6{,}000$ are shown in Fig. 6 and Fig. 7 respectively. The results indicate that the detectability is improved after the proposed method is employed. Specifically, with the proposed method for $k = 6{,}000$, the P_E of PSRM decreased by 19.68% for SUNIWARD with 0.4 bpp, 17.49% for SUNIWARD with 0.5 bpp. For SRM, the P_E decreased by 1.56% for WOW with 0.4 bpp, and 5.45% for WOW with 0.5 bpp. For the cases of $k = 5{,}000$, the P_E of PSRM decreased by 16.83% for SUNIWARD with 0.4 bpp, 15.6% for SUNIWARD with 0.5 bpp.

For some cases, the P_E has not decreased. This may be caused by the conditions of the image selection algorithm for producing re-training pseudo-labels. It is possible that neither sets of conditions are satisfied so no re-training is possible, therefore no improvement can be achieved. While no improvement in this case, the re-training process does not make the base-line classifier's performance worse.

In real world, it may be difficult to obtain 10,000 images for detection and the large ratio of stego images for a large image set is rare. To verify the practicability of the proposed method with smaller sets of images, we randomly chose 100 and 200 images from BOSSbass ver. 1.01 to form two small image sets ($n = 100$ and $n = 200$). The corresponding experimental results are shown in Fig. 8, where the steganographic method is SUNIWARD and the payload is 0.5 bpp. Since all the results of P_E are the average value of ten independent tests, the variances of P_E are also given in Table 1. It can be seen that the fluctuation of P_E is not large. From Fig. 8, we can see that the detection error rate also decreases for small image sets in most cases using the proposed method. Therefore, the proposed method can be effective for smaller image sets, which could allow closer inspection of one or a few target user(s) who are use steganography actively (i.e., with a large ratio of k/n).

We did not compare our method with the method in [35]–[39] because the two kinds of steganalytic methods work at different levels. The method in [35]–[39] aims to recognize the users of steganography while ours aims to recognize stego-images. In other words, our method focuses on detecting the existence of the secret information within images with the help of behavioral analysis (cover selection) of users. We however plan to extend our method in future to recognize users of steganography and then conduct a comparison of its performance with that of the work in [35]–[39].

TABLE 1. Variances ($\times 10^{-3}$) of detection errors for smaller image sets ($n = 100$ and $n = 200$).

	$n = 100$ PSRM / PSRM-P	$n = 200$ PSRM / PSRM-P
$k/n = 0.1$	0.644 / 1.261	1.560 / 1.536
$k/n = 0.2$	0.384 / 0.269	1.431 / 0.905
$k/n = 0.3$	0.256 / 0.089	0.595 / 0.464
$k/n = 0.4$	0.576 / 0.196	0.106 / 0.166
$k/n = 0.5$	0.536 / 0.681	0.130 / 0.251
$k/n = 0.6$	0.225 / 0.536	0.235 / 1.164
$k/n = 0.7$	0.465 / 0.721	0.377 / 0.845
$k/n = 0.8$	0.121 / 0.121	0.197 / 0.212
$k/n = 0.9$	2.581 / 2.581	0.316 / 0.250

V. CONCLUSION

This paper proposes a new steganalytic method using a clustering algorithm to improve the performance of a base-line classifier. The improvement is achieved by re-training a base-line classifier, which is pre-trained based on a standard image database, towards the actual cover images used by the user of the steganography. The re-training is done based on pseudo-labels of images under inspection, so the re-trained classifier is effectively more "contextualized" to perform better. Our experimental results proved the proposed method worked with even a smaller set of suspected images and a simple clustering algorithm like k-means.

REFERENCES

[1] J. Fridrich, *Steganography in Digital Media: Principles, Algorithms, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[2] G. Feng, X. Zhang, Y. Ren, Z. Qian, and S. Li, "Diversity-based cascade filters for JPEG steganalysis," *IEEE Trans. Circuits Syst. Video Technol.*, to be published.
[3] B. Li, Z. Li, S. Zhou, S. Tan, and X. Zhang, "New steganalytic features for spatial image steganography based on derivative filters and threshold LBP operator," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 5, pp. 1242–1257, May 2018.
[4] J. Kodovský, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 432–444, Apr. 2012.
[5] 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.
[6] 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.
[7] T. Denemark, V. Sedighi, V. Holub, R. Cogranne, and J. Fridrich, "Selection-channel-aware rich model for steganalysis of digital images," in *Proc. IEEE Int. Workshop Inf. Forensics Secur.*, Dec. 2014, pp. 48–53.
[8] J. Ni, J. Ye, and Y. I. Yang, "Deep learning hierarchical representations for image steganalysis," *IEEE Trans. Inf. Forensics Security*, vol. 12 no. 11, pp. 2545–2557, Nov. 2017.
[9] 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. 5, pp. 1200–1214, May 2018.

[10] 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 2018.
[11] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 920–935, Sep. 2011.
[12] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP J. Inf. Secur.*, vol. 2014, p. 1, Dec. 2014.
[13] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *Proc. IEEE Int. Workshop Inf. Forensics Secur.*, Dec. 2012, pp. 234–239.
[14] B. Li, M. Wang, J. Huang, and X. Li, "A new cost function for spatial image steganography," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 4206–4210.
[15] 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.
[16] L. Guo, J. Ni, W. Su, C. Tang, and Y.-Q. Shi, "Using statistical image model for JPEG steganography: Uniform embedding revisited," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 12, pp. 2669–2680, Dec. 2015.
[17] Z. Wang, X. Zhang, and Z. Yin, "Hybrid distortion function for JPEG steganography," *J. Electron. Imag.*, vol. 25, no. 5, 2016, Art. no. 050501.
[18] M. Kharrazi, H. T. Sencar, and N. Memon, "Cover selection for steganographic embedding," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2006, pp. 117–120.
[19] O. Evsutin, A. Kokurina, and R. Meshcheryakov, "Approach to the selection of the best cover image for information embedding in JPEG images based on the principles of the optimality," *J. Decis. Syst.*, vol. 27, no. S1, pp. 256–264, Apr. 2018.
[20] H. Sajedi and M. Jamzad, "Cover selection steganography method based on similarity of image blocks," in *Proc. IEEE 8th Int. Conf. Comput. Inf. Technol. Workshops*, Jul. 2008, pp. 379–384.
[21] H. Sajedi and M. Jamzad, "Using contourlet transform and cover selection for secure steganography," *Int. J. Inf. Secur.*, vol. 9, no. 5, pp. 337–352, Oct. 2010.
[22] M. S. Subhedar and V. H. Mankar, "Curvelet transform and cover selection for secure steganography," *Multimedia Tools Appl.*, vol. 77, no. 7, pp. 8115–8138, Apr. 2018.
[23] R.-E. Yang, Z.-W. Zheng, and W. Jin, "Cover selection for image steganography based on image characteristics," *J. Optoelectron. Laser*, vol. 25, pp. 764–768, Apr. 2014.
[24] S. Wu, Y. Liu, S. Zhong, and Y. Liu, "What makes the stego image undetectable?" in *Proc. 7th Int. Conf. Internet Multimedia Comput. Service*, 2015, Art. no. 47.
[25] Z. Wang, X. Zhang, and Z. Yin, "Joint cover-selection and payload-allocation by steganographic distortion optimization," *IEEE Signal Process. Lett.*, vol. 25, no. 10, pp. 1530–1534, Oct. 2018.
[26] I. Lubenko and A. D. Ker, "Steganalysis with mismatched covers: Do simple classifiers help?" in *Proc. 14th ACM Workshop Multimedia Secur.*, 2012, pp. 11–18.
[27] J. Kodovskỳ, V. Sedighi, and J. Fridrich, "Study of cover source mismatch in steganalysis and ways to mitigate its impact," *Proc. SPIE*, vol. 9028, Feb. 2014, Art. no. 90280J.
[28] J. Pasquet, S. Bringay, and M. Chaumont, "Steganalysis with cover-source mismatch and a small learning database," in *Proc. 22nd Eur. Signal Process. Conf.*, Sep. 2014, pp. 2425–2429.
[29] Z. Li and A. G. Bors, "Selection of robust features for the cover source mismatch problem in 3D steganalysis," in *Proc. 23rd Int. Conf. Pattern Recognit.*, Dec. 2016, pp. 4256–4261.
[30] S. Kouider, M. Chaumont, and W. Puech, "Technical points about adaptive steganography by oracle (ASO)," in *Proc. 20th Eur. Signal Process. Conf.*, Aug. 2012, pp. 1703–1707.
[31] Z. Wang and X. Zhang, "Secure cover selection for steganography," *IEEE Access*, vol. 7, pp. 57857–57867, 2019.
[32] X. Zhang, F. Peng, and M. Long, "Robust coverless image steganography based on DCT and LDA topic classification," *IEEE Trans. Multimedia*, vol. 20, no. 12, pp. 3223–3238, Dec. 2018.
[33] S. Li and X. Zhang, "Toward construction-based data hiding: From secrets to fingerprint images," *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1482–1497, Mar. 2019.
[34] H. Shi, J. Dong, W. Wang, Y. Qian, and X. Zhang, "SSGAN: Secure steganography based on generative adversarial networks," in *Proc. 18th Pacific-Rim Conf. Multimedia*. Harbin, China: Springer, Sep. 2017, pp. 534–544.
[35] A. D. Ker and T. Pevný, "A new paradigm for steganalysis via clustering," *Proc. SPIE*, vol. 7880, Feb. 2011, Art. no. 78800U.
[36] A. D. Ker and T. Pevný, "The steganographer is the outlier: Realistic large-scale steganalysis," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 9, pp. 1424–1435, Sep. 2014.
[37] T. Pevnỳ and I. Nikolaev, "Optimizing pooling function for pooled steganalysis," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Nov. 2015, pp. 1–6.
[38] F. Li, K. Wu, J. Lei, M. Wen, Z. Bi, and C. Gu, "Steganalysis over large-scale social networks with high-order joint features and clustering ensembles," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 2, pp. 344–357, Feb. 2016.
[39] Z. Wang, Z. Qian, and X. Zhang, "Single actor pooled steganalysis," in *Proc. Int. Conf. Genetic Evol. Comput.* Singapore: Springer, 2018, pp. 339–347.
[40] 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.
[41] P. Bas, T. Filler, and T. Pevný, "'Break our steganographic system': The ins and outs of organizing BOSS," in *Proc. 13th Int. Workshop Inf. Hiding (IH)*. Prague, Czech Republic: Springer, May 2011, pp. 59–70.
[42] G. Schaefer and M. Stich, "UCID: An uncompressed color image database," *Proc. SPIE*, vol. 5307, pp. 472–481, Jan. 2003.
[43] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, Jul. 2002.

ZICHI WANG received the B.S. degree in electronics and information engineering and the M.S. degree in signal and information processing from Shanghai University, China, in 2014 and 2017, respectively, where he is currently pursuing the Ph.D. degree. His research interests include steganography, steganalysis, and reversible data hiding. He has published about 30 articles in these areas.

SHUJUN LI (M'08–SM'12) received the B.E. degree in information science and engineering, in 1997, and the Ph.D. degree in information and communication engineering, in 2003, from Xi'an Jiaotong University, China. Since November 2017, he has been a Professor of cyber security with the University of Kent, U.K., leading the university wide Kent Interdisciplinary Research Centre in Cyber Security (KirCCS), a U.K. government recognized Academic Centre of Excellence in Cyber Security Research (ACE-CSR). He has published over 100 scientific articles with two Best Paper Awards. His research interests include cyber security, human–computer interface, multimedia computing, digital forensics, and cybercrime.

XINPENG ZHANG (M'11) received the B.S. degree in computational mathematics from Jilin University, China, in 1995, and the M.E. and Ph.D. degrees in communication and information system from Shanghai University, China, in 2001 and 2004, respectively. Since 2004, he has been with the faculty of the School of Communication and Information Engineering, Shanghai University, where he is currently a Professor. His research interests include information hiding, image processing, and digital forensics. He has published over 200 articles in these areas.