International Journal of Information Science and Management
Vol. 16, No. 1, 2018, 191-201

A Low Cost Image Steganalysis by Using Domain Adaptation

Mohammd Bagher Dastgheib
Assistant Prof. in Computer Engineering,
Research Department of Desinging and
Systemopration, RICeST, Iran.
Corresponding Author, dastgheib@ricest.ac.ir

Mahsa Farboudnia Jahromi
M.S in Azad University, Bushehr, Iran
mfarbod63@yahoo.com

Jafar Tahmoures Nejad
Assistant Prof. in Urmia University of
Technology Urmia, Iran
tahmores@gmail.com

Abstract

Information hiding and data encryption are used widely to protect data and information from anonymous access. In digital world, hiding and encrypting of the desired data into an image is a smart way to protect information with a low cost. In the digital images, steganalysis is a known method to distinguish between clean and stego images. Most of recent researches in this scope exploit feature reduction algorithms to improve the performance of correct detections. However, dimension reduction alone could not tackle the problem of steganalysis because the properties of stego images change during the steganalysis process. In this work, it is intended to propose an Image Steganalysis using visual Domain Adaptation (ISDA), which this steganalysis target images to distinguish across stego and clean images. ISDA is a dimensionality reduction approach that considers the image drifts during the steganography process in the steganalysis of target images. Moreover, ISDA employs domain invariant clustering in an embedded representation to cluster clean and stego images in the reduced subspace. The results on benchmark datasets demonstrate that ISDA thoroughly outperforms all of the state of the art methods on validation parameters, accuracy of detection and time complexity.

Keywords: Image Steganalysis, Visual Domain Adaptation, Feature Extraction, Embedded Representation.

Introduction

From a long time ago, people recognize the need to hide information via various ways. Thus, two closely related technologies invented to the steganography process, namely fingerprinting and watermarking. Steganography is a type of information hiding that means “covered writing”. In other words, steganography hides information beside regular information like pixels of an image.

Steganalysis is the science of studying and detecting messages that have been hidden using the steganography where it is an analogous term to cryptanalysis applied to cryptography. The goal of steganalysis is to identifying suspected packages, determining whether or not they have a payload encoded into them, and, if possible, recover that payload.

In fact, the most important information can be transferred through steganography techniques, which is a safe way that it cannot be attacked, detected or accessed.

Least significant bit (LSB) is a type of steganography in which the lowest bit plane of a bitmap image is used to convey the secret data (Akay and Karaboga, 2015; Ker, 2005; Xia et al., 2015,2016). The LSB method is used commonly because the eye cannot detect small perturbations embedded into an image. LSB methods for steganography are very simple to implement so most of free steganography tools uses this method (Akay et al., 2015; Miche et al., 2006). However, the detection of stego images from cover images is not very simple because in most of cases the original image is not available. Generally, steganalysis process is handled by a statistical analysis (Westfeld, 2001). Some simple methods use histogram or spectrum analysis for steganalysis process. In some cases, such as when only a single image is available, more complicated analysis techniques may be required. In total, steganography attempts to make distortion to the carrier indistinguishable from the carrier's noise floor. In practice, however, this is often improperly simplified to deciding to make the modifications to the carrier resemble white noise as closely as possible, rather than analyzing, modeling, and then consistently emulating the actual noise characteristics of the carrier (Akay et al., 2015;Carrier, 2011;Westfeld, 2001).

In general, each image could be categorized as a cover-image (clean image) or a stego-image (injected image). In the other words, images with no hidden message are called cover-image and images contain hidden message are called stego-image (Carrier, 2011;Ker, 2005). Steganalysis can be also considered as a pattern recognition process due to its similarity to the feature extraction methods. This process classifies an input as a stego or cover image. The features should be related positively or negatively to both of stego and clean images in order to distinguish them. The steganalysis is a complicated task and most of proposed methods cannot reach to a desired accuracy in the real world test cases (Akay et al., 2015; Bas and Fridrich, 2010).

Some of steganalysis approaches follow a conventional machine learning method, which consists of two steps. The first step extracts features from images, and the second step trains a standard classifier, e.g. SVM or FLD-based ensemble classifier, based on the extracted features (Liu, 2011). The major challenge in these methods lies in extracting effective representations to capture enough traces caused by embedding operations. Moreover, in the past decades, some researchers have focused on various handcrafted features. Although, significant progresses have been achieved in recent researches, the detection accuracy of current steganalysis systems based on handcrafted features is far from ideal results (Denemark et al., 2016). Moreover, the handcrafted feature designing is heavily dependent on expert experiences, and it is difficult and time-consuming to design new manual features.

In recent studies many researchers worked on steganalysis to improve detection performance (Denemark et al., 2016; Pevny et al., 2010). Most of proposed methods contain a feature reduction strategy to improve the detection accuracy. The most recent steganalysis approach exploited bee-colony beside feature selection algorithm and LSB to tackle steganalysis problem.

In this work we propose a novel feature extraction method to steganalysis the suspicious images. Image Steganalysis using visual Domain Adaptation (ISDA) stands on domain shift across images to detect stego-images. ISDA reduces joint marginal and conditional distributions across training and test sets (source and target domains, respectively) in an

unsupervised manner in an embedded subspace. Furthermore, ISDA benefit from condensed domain invariant clusters in the new representation to separate various classes of images. Moreover, ISDA adapt the image drifting produced by steganography to matching stego-images. ISDA shows stunning results on benchmark datasets against other available state of the art methods while standard classifiers often demonstrate poor recognitions due to significant difference across source and target domains.

The rest of paper is organized as follows. The next section contains a comprehensive literature review. Proposed method has been arranged in Section 3. Results and experiments are discussed in Section 4. The last section contains the conclusion and future works.

Related works

In recent years, many research studies used dimensionality reduction as a pre-analysis processing to separate the irrelevant and unimportant features from the relevant and important ones. The dimensionality reduction process is classified into feature selection and feature extraction methods. Feature selection methods are techniques of selecting a possible feature set from the whole set of candidate features. The later, namely, feature extraction method is a technique to extract necessity features from the original data (e.g. image) in order to reduce the dimension of input data. The feature selection methods can be considered as a branch of general feature extraction methods. In the rest, feature selection methods are reviewed briefly and then feature extraction methods that are the base of ISDA are explained by more details.

Unlike feature extraction methods, feature selection techniques have been applied to a set of data with identified features. The goal of this strategy is to remove irrelevant and redundant features and bold the important features in feature space. The feature selection method selects the optimal subset of features with the best performance that has less information loss. The feature selection methods, based on search and selection strategy, are categorized into three main categories: 1) complete 2) heuristic, and 3) stochastic (Pen and Yang, 2010;TahmoresNejad and Hashemi, 2016).

Flexible and robust heuristic feature selection approaches based on swarm intelligence algorithms are used in recent steganalysis researches. Mohammadi et al. (2014) proposed a novel approach to detect stego images based on bee colony feature selection method. The proposed approach selected stego-oriented features according to a heuristic to recognize stego- and cover images. Rostami et al. (2016) also used swarm optimization to improve steganalysis detection accuracy.

In recent years, many researchers exploited feature extraction and dimensionality reduction to distinguish stego- and cover images as well. In this way, they employed various vector sizes for features and also benefit from new features. Chen and Shi (2008) used Markov features using intra-block and inter-block dependencies. Kodovsky and Fridrich (2009) also enhanced Chen (2008) features using Cartesian calibration. Kodovsky and Fridrich (2011) proposed a high dimensional feature space. Bas et al. (2010) used subtractive pixel adjacency model for steganalysis. Kodovsky et al. (2012) proposed a compact rich model for Discrete Cosine Transform (DCT) domain and used this model for steganalysis. Pevny et al. (2007) also used a hybrid method that uses a combination of DCT and Markov features for multi-class JPEG steganalysis.

Christaline A. et al. (2016) proposed a metaheuristic approach based on random behavior of plants and animals. The proposed approach employed AntLion behavior based

Optimization technique (ALO) beside movement of ants. The model used random walk and the traps built by antlions. SVM, MLP and the fusion classifiers - Bayes, Decision template and Dempster Schafer are used to classify target images.

Qian et al. (2016) used Conventional Neural Network (CNN) to tackle the steganalysis problem. In this way, they used Transfer Learning (TL) to learn a CNN. The extracted feature representations with a pre-trained CNN employed to detect steganographic images with high payload.

In this paper, we propose a joint marginal and conditional distribution adaptation method that employs domain invariant clustering to discriminate between various images. ISDA transfers knowledge from the source to target domain by preserving statistical and geometric structure of domains in the embedded representation. Moreover, ISDA constructs condensed clusters in the embedded representation that are domain invariant and discriminative for target image data classification.

Proposed method

In this section, ISDA approach for effectively tackling the problem of steganalysis is presented in detail.

Motivation

Most of the conventional solutions for the problem of steganalysis benefit from the dimensionality reduction either feature selection or feature extraction without considering that the nature and the properties of images have been changed during the steganography process. However, the distribution of images before and after steganography procedure has significantly drifts. Thus, the reduced feature set from one domain (i.e. the image set before steganography process) will have considerable difference with another domain (i.e. the image set after steganography process). Then, the performance of trained model on the reduced source domain will degrade dramatically on the reduced target domain. Figure 1 demonstrates the distribution of histogram gradient energy (HGE) of stego- and cover images.

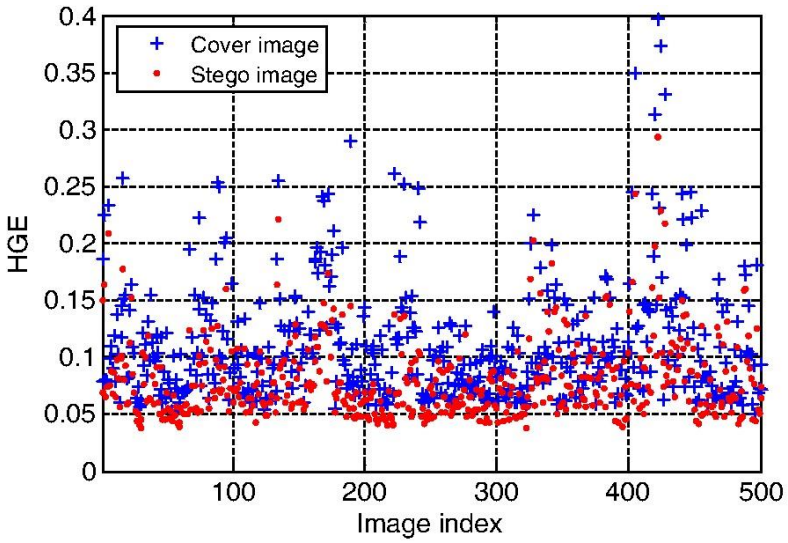

Figure 1. (Best viewed in color) Distribution of histogram gradient energy (HGE) of 500 stego- and cover images (Boss-base dataset[1]). As is clear, the distribution of stego- (mentioned with blue plus) and cover (mentioned by red circle) images have considerable difference.

However, we reduce the dimension of input data considering the following contributions. (1) We suppose that we are given an m-dimensional representation of data from $X_s \epsilon R^{m \times n_s}$ and$X_t \epsilon R^{m \times n_t}$, source and target domains with n_s and n_t samples respectively, and (2) we find a domain invariant representation across source and target domains so that adapt the distribution of stego- and cover images in the embedded representation.

Dimensionality reduction and domain adaptation

In this work, we choose Principal Component Analysis (PCA) as a baseline dimensionality reduction approach. The main goal of PCA is to find an intermediate representation which orthogonally transforms input matrix $X \epsilon R^{m \times n}$ with m features and n samples, into an embedded subspace with maximum variance via covariance matrix, XHX^T. $H = I - \frac{1}{n}1$ is centering matrix where I denotes identity matrix and 1 is the ones matrix. The transformation matrix $A \epsilon R^{m \times k}$ is achieved from $max_{A^T A = I} tr(A^T XHX^T A)$ on which $A^T A = I$ is orthogonality constraint, and tr denotes the trace of matrix (Pen and Yang, 2010; TahmoresNejad et al., 2016).

Most of traditional steganalysis approaches benefit from dimensionality reduction methods such as PCA. However, the distribution difference between stego- and cover images will still be considerable large in the embedded k-dimensional representation. Thus we employ a distance measure to compute distribution difference across source and target domains. Maximum Mean Discrepancy (MMD) is exploited as a non-parametric metric to compute distribution difference across domains. MMD computes the distance between the sample means of source and target sets in the k-dimensional Reproducing Kernel Hilbert Space (RKHS). The following relation demonstrates the MMD:

$$MMD(X_s . X_t) = \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \varphi(x_i^s) - \frac{1}{n_t} \sum_{i=1}^{n_t} \varphi(x_i^t) \right\|_H^2 \quad (1)$$

Where φ is the feature map defined as $\varphi(x): X \to H$, and H denotes a universal RKHS. The Equation 1 could be considered as $tr(XW_0X^T)$ in closed form, where $W_0 \epsilon R^{(n_s+n_t)\times(n_s+n_t)}$ is a composite MMD matrix in the form $\begin{bmatrix} W_0^{ss} & W_0^{st} \\ W_0^{ts} & W_0^{tt} \end{bmatrix}$ and $W_0^{ss} = \frac{1}{n_s n_s}$, $W_0^{tt} = \frac{1}{n_t n_t}$ and $W_0^{st} = \frac{-1}{n_s n_t}$ are source, target and cross domains MMD matrices. tr denotes the trace of matrix and $X = \{X_s . X_t\}$ (Molina et al., 2002; TahmoresNejad et al., 2016)

To reduce the difference between marginal distributions $P_s(x_s)$ and $P_t(x_t)$, we adapt MMD as the distance measure to minimize distribution difference across source and target domains. Let $A \epsilon R^{m \times k}$ denote the transformation matrix, where transforms source and target data into an embedded subspace with minimum distance. Thus the objective function to minimize marginal distribution difference will be as follows:

$$Margial(X_s . X_t) = tr(A^T XW_0 X^T A) \quad (2)$$

[1] Please see home page: http://agents.fel.cvut.cz/boss/index.php?mode=VIEW&tmpl=home

However, steganography intensifies conditional distribution difference across stego- and cover images in addition to marginal distribution difference. Thus we are to minimize conditional distribution difference between source and target domains. Here we customize empirical MMD to measure the distance between the class-conditional distributions.

$$MMD_{cond}(X_s.X_t) = \left\| \frac{1}{n_c^s} \sum_{x_i \epsilon X_c^s} \varphi(x_i) - \frac{1}{n_c^t} \sum_{x_i \epsilon X_c^t} \varphi(x_i) \right\|_H^2 \quad (3)$$

where n_c^s and n_c^t denote the number of source and target samples that belongs to class c, respectively. Also, X_c^s and X_c^t are the source and target samples from class$c$, respectively. The Equation 3 could be considered as $tr(XW_cX^T)$ in closed form, where $W_c \epsilon R^{(n_s+n_t)\times(n_s+n_t)}$ is a composite MMD matrix in the form $\begin{bmatrix} W_c^{ss} & W_c^{st} \\ W_c^{ts} & W_c^{tt} \end{bmatrix}$ and $W_c^{ss} = \frac{1}{n_c^s n_c^s}$, $W_c^{tt} = \frac{1}{n_c^t n_c^t}$ and $W_c^{st} = \frac{-1}{n_c^s n_c^t}$ are source, target and cross domains MMD matrices.

Since the target domain is unsupervised, the values of n_c^t for various classes are unknown. In this way, we employ source data to build a model for target data label prediction. It's clear that the predicted labels are imprecise; however, they could be exploited to calculate W_c in an iterative manner (Molina et al., 2002).

To reduce the difference between conditional distributions $P_s(x_s|y_s = c)$ and $P_t(x_t|y_t = c)$, we adapt MMD as the distance measure to minimize distribution difference across source and target domains. Thus the objective function to minimize the conditional distribution difference will be as follows:

$$Conditional(X_s.X_t) = tr(A^T X W_c X^T A) \quad (4)$$

Moreover, ISDA benefit from domain invariant clustering to minimize within-class scatter across stego- and cover images. In this way, ISDA minimizes the distance of each transformed source sample from its projected mean. Thus the following relation is minimized where μ^c denotes the mean of class c.

$$tr(A^T \sum_{\forall c \epsilon C} \sum_{x_i \epsilon X_c^s} (x_i - \mu^c)^T (x_i - \mu^c) A) \quad (5)$$

In ISDA, to find an effective and robust transformation, we simultaneously minimize the marginal and conditional distribution differences and also, within-class scatter matrix. Thus the objective function is composed from Equations 2, 4 and 5 on PCA optimization problem (TahmoresNejad et al., 2016).

$$min_{A^T X H X^T A = I} \left(tr(A^T X W_0 X^T A) + tr(A^T X W_c X^T A) + tr\left(A^T \sum_{\forall c \epsilon C} \sum_{x_i \epsilon X_c^s} (x_i - \mu^c)^T (x_i - \mu^c) A\right) + \gamma \|A\|_F^2 \right) \quad (6)$$

where $\|\cdot\|_F^2$ denotes the Frobenius norm of transformation matrix A to avoid from trivial solutions. To solve the Equation 6, we derive the Lagrange function and differentiate according to transformation parameter A. Thus the generalized eigen-decomposition is achieved as follows:

$$XW_0X^TA + XW_cX^TA + \sum_{\forall c \epsilon C} \sum_{x_i \epsilon X_c^s} (x_i - \mu^c)^T (x_i - \mu^c) A = XHX^TA\delta \quad (7)$$

Where δ is the Lagrange multiplier. The adaptation matrix is achieved from k smallest eigenvectors of Equation 7.

ISDA benefit from an iterative procedure to predict the target labels. However, we can usually obtain more accurate labelling in each iteration. This procedure is an Expected-Maximization (EM-) like process that refines results in each iteration. The refinement procedure is shown in the next section.

Experiments and Results

To reveal the performance of the proposed method and comparing it with other well-known methods, a set of experiments has been set up. One of the most important aspects of any performance evaluation is to use a standard data set with a variety of image textures. The proposed scheme employs the image database of BOSS version 1.01 that consists of 10,000 gray-scale images sized 512×512 pixels which is also used in modern steganographic schemes with embedding rates less than or equal to 1 bpp. So the BOSS dataset used here has 10,000 clean images as same as stego images. The BOSS-base dataset is used to evaluate the steganalysis in the literature. The proposed method was implemented and executed using MATLAB R2012a on an Intel Core i5-2500, 3.3 – 3.6GHz, with 8 GB RAM.

Extracting feature from images

In this paper both of subtractive pixel adjacency model (SPAM) method and Cartesian-calibrated PEV (CC-PEV) features are employed to extract the final feature set for steganalysis. The SPAM method has 686 features and the CC-PEV method has 548 features, so the final feature set for each image has 686+548=1,194 features. As sown in the fig.1 this is the first step of the proposed scheme. At the end of this step each image vector has 1,194 elements.

Algorithm 1 Visual domain adaptation (VDA)

Input: source and target data X; source domain labels y_s; regularization parameter λ; #subspace bases k
Output: target domain labels y_t
$$(W_0)_{ij} = \begin{cases} \frac{1}{n_s n_s} & \text{if } x_i, x_j \in \mathcal{D}_s \\ \frac{1}{n_t n_t} & \text{if } x_i, x_j \in \mathcal{D}_t \\ \frac{-1}{n_s n_t} & \text{otherwise} \end{cases}$$
$$S_w = (x_{i*}^{s^c} - \mu^c)(x_{i*}^{s^c} - \mu^c)^T \quad \forall i = 1 \ldots n_s,\ c = 1 \ldots C$$
$v = ones(n, n)$; an $n \times n$ matrix of ones
$H = I - \frac{1}{n} vv^T$
repeat until convergence
solve eigendecomposition $(X \sum_{c=0}^{C} W_c X^T + S_w + \lambda I)A = XHX^T A\phi$ and select k smallest eigenvectors as adaptation matrix A
update pseudo target labels using a standard classifier f trained on projected source data $\{A^T X_s, y_s\}$
$$\text{update } (W_c)_{ij} = \begin{cases} \frac{1}{n_s^c n_s^c} & \text{if } x_i, x_j \in \mathcal{D}_s^c \\ \frac{1}{n_t^c n_t^c} & \text{if } x_i, x_j \in \mathcal{D}_t^c \\ \frac{-1}{n_s^c n_t^c} & \text{if } x_i \in \mathcal{D}_s^c, x_j \in \mathcal{D}_t^c \parallel x_j \in \mathcal{D}_s^c, x_i \in \mathcal{D}_t^c \\ 0 & \text{otherwise} \end{cases}$$
end repeat
return target domain labels y_t determined by classifier f

Figure 2. The VDA algorithm used to extract the most effective features

Feature reduction using VDA

The different feature sets can achieve different accuracy. So it is common that select a subset of features to maximize the accuracy. In this work VDA is used to select the optimal feature set. The main idea of VDA is that embeds source and target data into a latent space on which minimizes marginal and conditional distribution differences and cluster same label instances. VDA procedure is an iterative process that converges based on increasing amount of true labels. As shown in figure 2, in each iteration VDA exploits pseudo-labeling besides optimization problem (EM-like) to refine the predicted labels. In general, VDA finds the labels of target data in an iterative manner.

Method evaluation

To compare our steganalysis results to other works, we set up two set of experiments. The first set of experiments compares the results of extracted features without any feature reduction. In all of these experiments we use K-NN classifier for steganalysis process. As seen in Fig 2, the accuracy of K-NN classifier is compared for true positive detection of stego images. In the Fig 2, the SPAM features, CC-PEV features and a mix of these features are used to set up the experiment. The result showed (Fig 2) that there is no valuable change in accuracy of detection stego images.

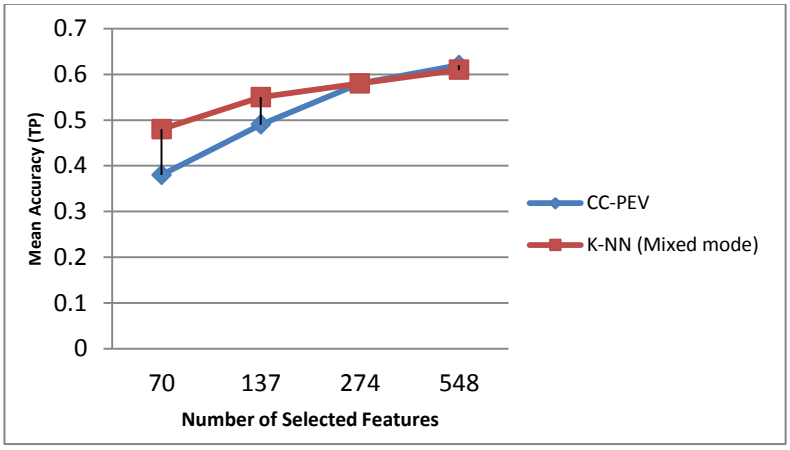

Figure 2. The Result of CC-PEV vs K-NN (mix of CC-PEV and SPAM features) without using VDA

In the second set of experiments, the results of feature selection based on bee colony are compared with result of feature extraction based on VDA. The accuracy of different feature sets in these schemes is compared to evaluate the proposed method.

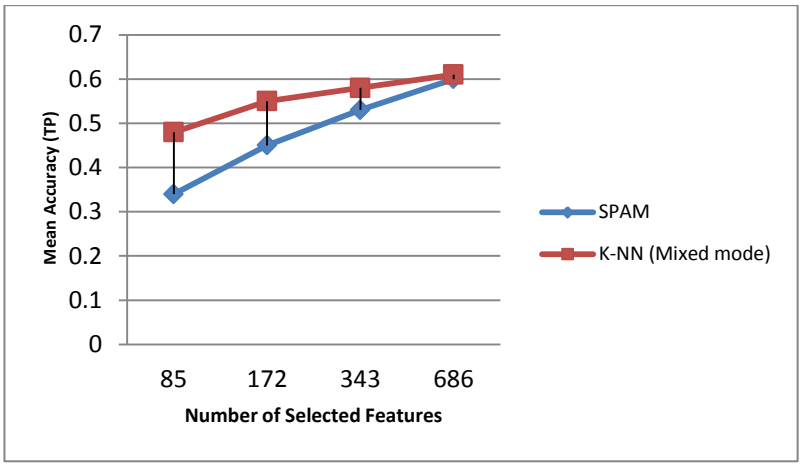

Figure 3. The Result of SPAM vs K-NN (mix of CC-PEV and SPAM features) without using VDA

To compare the proposed scheme, we used VDA to extract the best effective 20, 50, 100 and 200 features. The results of these experiments are showed in table 1. Fig 4 shows the result of the proposed VDA method with other methods. The proposed method can reach to 83% of accuracy in detection of stego images. The IFAB method is compared with the proposed method. The accuracy of VDA method is better than IFAB (Fig 4). The computational complexity of VDA algorithm is so lighter than IFAB that is based on bee colony algorithm.

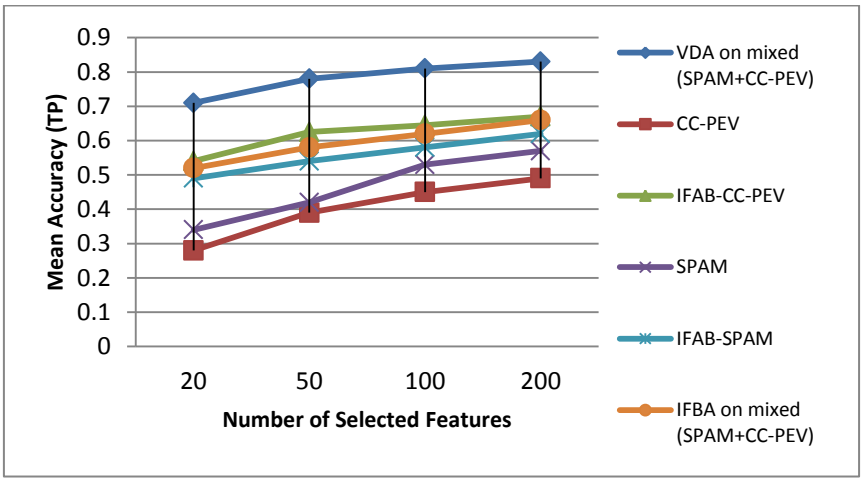

Figure 4. The result of VDA in comparison to some steganalysis methods

Table 1. The Results of proposed scheme (VDA)

No of features	20	50	100	200
KNN Accuracy	71%	78%	81%	83%

The result of table 1 is addressed by changing the value of K in K-NN classifier to get the best result. The average value of K for the best result is 10.

Conclusion

In this work, a novel feature extraction method based on Visual Domain Adaptation (VDA) is proposed to extract the optimal feature subset for steganalysis. To compare the

results of proposed scheme with other well-known methods like IFAB, the feature extracted from SPAM and CC-PEV and a combination of these features are used for evaluation. The result of evaluation showed that using these effective feature extraction methods without feature reduction can achieve poor accuracy. The results of VDA feature extraction method in comparison to the IFAB method can reach to better accuracy in a low cost polynomial time. The proposed method out performs the IFAB method in both accuracy and time complexity evaluations. The proposed feature extraction method based on VDA is completely effective and can be used in other domain of noisy image processing like OCR.

References

Akay, B., & Karaboga., D. (2015). A survey on the applications of artificial bee colony in signal, image, and video processing, *Signal, Image Video Process.*, 9(4), 967–90.

Anita Christaline, J., Ramesh, R., & Vaishali, D. (2016). Bio-Inspired computational algorithms for Improved Image Steganalysis, *Indian Journal of Science and Technology.*, 9(10).

Bas, P., & Fridrich, J. (2010). Steganalysis by Subtractive Pixel Adjacency Matrix, *Distribution*, 5(2), 215–24.

Carrier, C. (2011). A survey of steganography and steganalysis technique in image, text, audio and video as cover. *International Journal of Global Research in Computer Science*, 2(4). 36-46.

Chen, C., & Shi, Y. Q. (2008). JPEG image steganalysis utilizing both intrablock and interblock correlations, in *IEEE International Symposium on Circuits and Systems*, 3029–3032.

Denemark, T., Member, S., Boroumand, M., & Member, S. (2016). Steganalysis features for content-adaptive JPEG steganography. *IEEE Transactions on Information Forensics and Security*, 11(8), 1736-1746.

Fernando, B., Habrard, A., Sebban, M., & Tuytelaars, T. (2013). Unsupervised visual domain adaptation using subspace alignment, in *Proceedings of the IEEE International Conference on Computer Vision*, 2960–2967.

Ker, A. D. (2005). Steganalysis of LSB matching in grayscale images. *IEEE signal processing letters*, 12(6), 441-444.

Kodovsk`y, J., & Fridrich, J. (2009). *Calibration revisited. In Proceedings of the 11th ACM workshop on Multimedia and security* (pp. 63-74). ACM.

Kodovský , J. & Fridrich, J. (2011) Steganalysis in high dimensions: Fusing classifiers built on random subspaces. *In Media Watermarking, Security, and Forensics III*(Vol. 7880, p. 78800L).

Kodovský, J., & Fridrich, J. (2012). Ensemble Classi fi ers for Steganalysis of Digital Media, 7(2), 432–444.

Liu, Q. (2011). Steganalysis of DCT-embedding based adaptive steganography and YASS, in *Proceedings of the thirteenth ACM multimedia workshop on Multimedia and security*, 77–86.

Miche, Y., Roue, B., Lendasse, A., & Bas, P. (2006). A feature selection methodology for steganalysis, in *International Workshop on Multimedia Content Representation, Classification and Security*, 49–56.

Mohammadi, F. G. , & Abadeh, M. S. (2014). Image steganalysis using a bee colony based

feature selection algorithm. *Engineering Applications of Artificial Intelligence*, 31, 35-43.

Molina, L. C., Belanche, L., & Nebot, À. (2002). Feature selection algorithms: A survey and experimental evaluation. In Data Mining, 2002. ICDM 2003. Proceedings. *2002 IEEE International Conference on* (pp. 306-313). IEEE.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.

Pevny, T., Bas, P., & Fridrich, J. (2010). Steganalysis by subtractive pixel adjacency matrix. *IEEE Transactions on information Forensics and Security*, 5(2), 215-224.

Pevny, T., & Fridrich, J. (2007). Merging Markov and DCT features for multi-class JPEG steganalysis. *In Security, Steganography, and Watermarking of Multimedia Contents IX*. 65:(5), 62.

Qian, Y., Dong, J., Wang, W., & Tan, T. (2016). Learning and transferring representations for image steganalysis using convolutional neural network, in *Image Processing (ICIP), 2016 IEEE International Conference on*, 2752–56.

Rostami, V., & Khiavi, A. S. (2016). Particle Swarm Optimization based feature selection with novel fitness function for image steganalysis, in *Artificial Intelligence and Robotics (IRANOPEN)*, 109–14.

Tahmoresnezhad, J., & Hashemi, S. (2016). Visual domain adaptation via transfer feature learning. *Knowledge and Information Systems*, 50(2), 585-605.

Westfeld, A. (2001). F5 A Steganographic Algorithm, in *Information Hiding: 4th International Workshop, IH 2001, Pittsburgh, PA, USA, April 25-27, 2001. Proceedings*, 2137, 289.

Xia, Z., Wang, X., & Sun, X. (2014). Steganalysis of LSB matching using differences between nonadjacent pixels. *Multimedia Tools and Applications.* 75 (4), 1947-1962.

Xia, Z., Wang, X., Sun, X., Liu, Q., & Xiong, N. (2016). Steganalysis of LSB matching using differences between nonadjacent pixels, *Multimed. Tools Appl.*, 75(4), 1947–1962.