Visnyk NTUU KPI Seriia – Radiotekhnika Radioaparatobuduvannia, 2022, Iss. 89, pp. 54–60

UDC 621.39

Investigation of Digital Image Preprocessing Methods Influence on the Accuracy of Stego Images Detection

Progonov D. O.

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

E-mail: *progonov@gmail.com*

The feature of modern methods of detecting unauthorized transmission of confidential data in communication systems is widespread usage of pre-processing methods for transmitted files, such as digital images. The purpose of these methods is to detect weak changes of cover image’s statistical parameters caused by message hiding. A significant number of these methods are based on usage of ensembles of high-pass filters, which allows to ensure high accuracy of detection of steganograms (more than 95%) formed according to known steganographic methods. However, a significant limitation of the practical application of these methods is high computational complexity of ensemble forming procedure that minimizes the detection error of stego images. This makes it impossible to quickly reconfigure stegdetectors to detect stego images formed according to a priori unknown embedding methods. Therefore, it is of special interest to develop fast methods for image pre-processing, which can reliably detect weak changes of cover’s statistical parameters under limited a priori information about used steganographic method. The work is devoted to the study of the achievable accuracy of the stedetector with variations type and parameters of digital images pre-processing methods. According to the results of the study, the optimal methods of pre-processing image to minimize the detection error of stego images are proposed. These methods can significantly (up to 9 times) reduce the error of stego images detection compared to modern pre-processing methods, even in the most difficult case of low payload of cover image (less than 10%) and limited a priori data about used embedding method. It is revealed that usage of special types of image pre-processing methods, namely denoising autoencoders, allows to bring the accuracy of a stegdetector closer to the proposed estimations of achievable accuracy of stegodetectors.

Keywords: steganalysis; stego image preprocessing methods; digital images

DOI: 10.20535/RADAP.2022.89.54-60

Acronyms

ANN Artificial Neural Network

CI Cover Image

CNN Convolutional Neural Network

DAE Denoising Autoencoder

DI Digital Image

HPF High-Pass Filter

PPM Preprocessing Methods

SD Stegdetector

SM Steganographic Method

Introduction

Ensuring of reliable protection of sensitive data from unauthorized transmission (leakage) is topical task today. Of special interest are methods related to revealing and counteraction of hidden (concealed) data transmission with usage of multimedia data, such as Digital Image (DI) [1]. These methods are based on recent advances of digital media steganography, namely adaptive message hiding in DI, that makes revealing of formed stego images a non-trivial task.

The state-of-the-art approach to detect of stego images formed by novel Steganographic Method (SM) is based on analysis differences between the statistical, spectral, and structural parameters of initial (cover) and processed images [1]. This makes possible usage of multidimensional vector classification methods to reveal weak alterations of Cover Image (CI) parameters caused by stegodata embedding. However, the detection accuracy of novel Stegdetector (SD) hardly depends on the availability of prior information about used embedding methods that limits usage of SD in real cases where such information is limited or even absent.

One of proposed approaches to overcome mentioned limitation of modern SD is based on applying of Preprocessing Methods (PPM) to analyzed DI. The pre-processing methods are aimed at revealing and emphasizing weak alterations of CI pixels brightness caused by message hiding. Variety of proposed PPM makes possible fine-tuning of a SD to achieve high detection accuracy (more than 90%) for wide range of modern embedding methods. However, there is no information in the literature about the achievable detection accuracy for stego images by using PPM of specific type. This makes it difficult to choose effective pre-processing methods of DI to minimize the error of stego images detection especially for unknown embedding methods (zero-day problem). Thus, the topical task is estimation of achievable accuracy of stego images detection in case of limited or even absent prior information about used SM.

1 Related works

The considerable amount of proposed PPM for modern stegdetectors is based on usage an ensemble of High-Pass Filter (HPF) [1]. Applying of such filters allows effectively suppression of cover image content and further extraction of high-frequency components used for message hiding. Despite high accuracy of stego images detection (more than 95%) by usage of such approach to DI preprocessing, practical use of the corresponding SD is limited. This is caused by necessity of time-consuming and resource-intensive optimization of ensemble to minimize the stego images detection error.

To overcome this limitation, it was proposed to use Artificial Neural Network (ANN), in particular Convolutional Neural Network (CNN) and Denoising Autoencoder (DAE). The feature of ANN-based stegdetectors is ability to adapt neural networks parameters during training (backpropagation procedure) to minimize stego images detection error P_E [2]. This allows selecting the optimal parameters of input (convolutional) layers by criterion of error P_E minimization during backpropagation procedure instead of painstaking selection of HPF ensembles elements. In contrast to CNN, the DAE is based on projection of multidimensional vectors (related to statistical parameters of analyzed images) to a space of smaller dimensions while preserving their relative positions [2]. This makes it possible to reduce of P_E value on new datasets of DI in comparison with the widespread stegdetectors [5]. Despite outstanding detection accuracy of SD based on considered CNN and DAE architectures, tuning of such stegdetectors remain resource-intensive procedure. Also, trained stegdetectors may remain vulnerable to change of used image dataset (domain mismatch problem) due to usage of shallow neural networks caused by computation resource constraints.

Therefore, of special interest are development of advanced PPM methods that provide high detection accuracy of CI alteration caused by message hiding by preserving fixed (low) computation complexity. Despite variety of proposed PPM methods, selection of appropriate image pre-processing method that minimizes P_E under specified SM and used image dataset requires exhaustive search among known PPM methods [6]. This is caused by absence of theoretical foundations for the selection of optimal PPM based on the criterion of minimizing the value of the classification error of stego images P_E. Solving of mentioned problem requires estimation achievable detection accuracy depends on available information about used embedding method and digital image dataset. However, information presented in the literature covers only specific cases, such as standard image datasets or widespread SM. Thus, the work is aimed at filling this gap by analysis of achievable P_E values level by using modern PPM depending on the available priori information about used SM.

2 Task and challenges

The purpose of the work is estimation of achievable detection accuracy of stego images by varying the type of image pre-processing methods and limitations of priori data about SM. The results of research allows developing the methodology for determining the optimal PPM method by criterion of minimizing stego images detection error for modern steganographic methods.

3 Analysis of the achievable detection accuracy for moderd stegdetectors

The detection accuracy of the pre-trained SD can be represented in the following form, depending on the selected PPM and methods for DI features extraction:

$$P_E(\mathbf{I}) = f(\mathsf{F}, \mathsf{S}, \mathsf{K}, \mathbf{I}, \mathcal{X}, \mathcal{Y}), \qquad (1)$$

where P_E is stego images detection error; $\mathbf{I}$ is analyzed digital image; F is operator of image preprocessing; S is operator of feature extraction for processed DI; K is operator of comparison calculated parameters of the analyzed image with corresponding characteristics of the cover and stego images; $\mathcal{X}, \mathcal{Y}$ are sets of cover and stego images respectively that are available during SD training.

The operator K in eq. (1) corresponds to applying of PPM to analyzed image in order to detect weak changes of its statistical parameters cause by message embedding. Currently, the ensembles of HPF are widely used as such operator [3,6]. At the next stage, the operator S is applied to the extracted high-frequency components of processed analyzed image for estimation their statistical, spectral, and structural parameters [1]. However, ensuring high accuracy of stego images detection (more than 95%) requires usage of large number of high-pass filters (e.g. 12,870 parameters for the PSRM model, 34,671 for the maxSRM model), which complicates tuning of the SD. Finally, extracted features of pre-processed image is used for operator K in eq. (1). The ensemble classifiers, such as random Forests, are widely used for performing this task to minimize the value of the error P_E while preserving low computational complexity of the SD tuning [8].

Despite the emergence of advanced methods for analysis and classification of DI statistical parameters for design of high-precision SD, the selection of optimal PPM to minimize the P_E values currently is not given enough attention by researches. The common way to compare the effectiveness of PPM methods is based on investigation of shift for multidimensional vectors (statistical parameters) of cover and stego images caused by image pre-processing [7]. A schematic representation of the changes in the mutual position of these feature vectors for both cover and stego images after the applying of PPM is shown at Fig. 1.

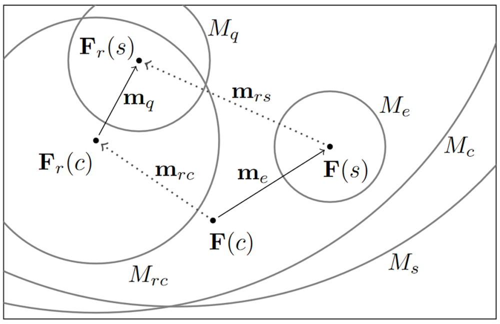

Fig. 1. The influence of digital image preprocessing methods on the mutual positions of vectors (statistical characteristics) of cover and stego images. According to materials [7].

Message embedding to the CI leads to shift of the corresponding feature vectors (statistical parameters) of the cover images $\mathbf{F}(c)$ by the amount $\mathbf{m}_e$ to the new position $\mathbf{F}(s)$. In this case, the spread of the values of the vectors $\mathbf{F}(c)$ is denoted as $\mathbf{M}_c$, while the corresponding spread of the vectors $\mathbf{F}(s)$ for formed stego images is equal to $\mathbf{M}_e$. Applying of PPM methods leads to a shift of the vectors $\mathbf{F}(c)$ and $\mathbf{F}(s)$ by the corresponding values $\mathbf{m}_{rc}$ and $\mathbf{m}_{rs}$ to the new positions $\mathbf{F}_r(c)$ and $\mathbf{F}_r(s)$. Reciprocally, the value of the "spread" of positions for feature vectors related to processed cover and stego images equals to $\mathbf{M}_{rc}$ and $\mathbf{M}_q$ respectively.

Applying of PPM to the studied DI leads to corresponding changes of mutual positions for clusters of extracted feature vectors of cover and stego images. In the work [7], the following classification of DI preprocessing methods was proposed:

1. Parallel reference — application of PPM leads to the similar changes of characteristics for cover $\mathbf{F}_r(c)$ and stego $\mathbf{F}_r(s)$ images;

2. Eraser — application of these methods leads to a decrease in the distance between $\mathbf{F}_r(c)$ and $\mathbf{F}_r(s)$ compared to the distance between the vectors $\mathbf{F}(c)$ and $\mathbf{F}(s)$;

3. Divergent reference — aimed at increasing the distance between clusters of vectors $\mathbf{F}_r(c)$ and $\mathbf{F}_r(s)$;

4. Cover estimate — is aimed at evaluating the statistical features of the CI based on the available (noisy) images;

5. Stego estimate — is aimed at detecting distortions of the CI caused by hiding the recorded data.

Let us note that usage of PPM methods related to parallel reference or eraser cases is limited today. This is caused by negligible impact on distance between the clusters of vectors $\mathbf{F}_r(c)$ and $\mathbf{F}_r(s)$ and, respectively, detection accuracy. On the other hand, divergent reference methods allow considereably improving detection accuracy by strengthening the differences between the cover and stego images. However, practical usage of such methods requires precise assessment relative position of feature vectors clusters $\mathbf{F}_r(c)$ and $\mathbf{F}_r(s)$. This needs usage of prior information about features of used embedding method that may be limited or even absent in real cases.

As a result, a significant number of modern PPM are aimed at estimating the parameters of the CI based on the available (noisy) data (cover estimate methods), or detecting weak changes of CI parameters caused by message hiding (stego estimate methods) [9, 10]. However, there is limited information about achievable accuracy of SD by using these types of PPM.

Let us consider the case of applying idealized PPM methods related to the cover $(\mathsf{F}^{CE}_{ideal}(\cdot))$ and stego $(\mathsf{F}^{SE)}_{ideal}(\cdot))$ estimation cases:

$$\mathsf{F}^{CE}_{ideal}(\mathbf{X},\mathbf{Y}) : \mathbf{X} \to \mathbf{X}, \mathbf{Y}\left(\Delta^S_\alpha\right) \xrightarrow[\forall \Delta^S_\alpha > 0]{} \mathbf{X}, \qquad (2)$$

$$\mathsf{F}^{SE}_{ideal}(\mathbf{X},\mathbf{Y}) : \mathbf{X} \xrightarrow[\forall \Delta^S_\alpha > 0]{} \mathbf{Y}\left(\Delta^S_\alpha\right), \quad \mathbf{Y}\left(\Delta^S_\alpha\right) \xrightarrow[\Delta^S_\alpha = const]{} \mathbf{Y}\left(\Delta^S_\alpha\right), \qquad (3)$$

where $\mathbf{X}, \mathbf{Y}$ are cover and stego images respectively; Δ^S_α is the cover image payload.

Let us denote the statistical parameters for original and processed images by the corresponding vectors $\mathbf{F}_{nc}$ and $\mathbf{F}_{calib}$. Then, applying of considered methods in eqs. (2) and (3) leads to the case when either features of cover, or stego images are preserved after transformation while counterpart features are changed considerably. For example, the elements of $\mathbf{F}_{calib}$ vector will be the same for cover (eq. (2)) or stego (eq. (3)) images, which will significantly reduce the accuracy of the SD. In this case, vectors $\mathbf{F}_{CC} = \{\mathbf{F}_{nc}; \mathbf{F}_{calib}\}$ will include the features of original and processed images, which is widely used in modern SD. On the other hand, the magnitude of vectors $\mathbf{F}_{DF} = \mathbf{F}_{nc} - \mathbf{F}_{calib}$ will be proportional to the CI features distortions caused by stego data embedding.

Then, accuracy of estimation of CI distortions caused by message hiding by using vectors $\mathbf{F}_{DF}$ depends on the accuracy of restoration cover image from current (noisy) ones (by using the operator $\mathsf{F}^{CE}_{ideal}(\mathbf{X},\mathbf{Y})$), or stego one (by using the operator $\mathsf{F}^{SE}_{ideal}(\mathbf{X},\mathbf{Y})$). Respectively, usage of "idealized" operators (eqs. (2) and (3)) will allow minimizing the stego image detection error for fixed methods for images features estimation and classification. Thus, applying of vectors $\mathbf{F}_{DF}$ makes possible estimation of achievable level of P_E error by using standard methods for DI feature extraction and classification.

4 Results

4.1 Experimental setup

Performance analysis of achievable accuracy of stego images detection by usage of proposed DI preprocessing methods (eqs. (2) and (3)) was carried out on the standard image package ALASKA [11]. The case of applying modern steganographic methods HUGO [12] and MiPOD [13] for the message embedding into CI was considered. These methods are based on presentation of embedding process as solving of corresponding optimization problem — minimizing changes of CI statistical parameters during embedding a message with fixed bitlength. The CI payload values were varied in the following range — 3%, 5%, 10%, 20%, 30%, 40%, 50%.

The analysis of detection accuracy was perfomed according to standard cross-validation procedure. The used images dataset was divided into training $\mathcal{S}_{train}$ (70%) and $\mathcal{S}_{test}$ testing (30%) subsets. The partitioning was repeated 10 times to obtain the average P_E values. Images from the training and testing sets were scaled to the identical resolution of 512×512 pixels.

The tuning of SD was done by varying the ratio of cover and stego images pairs (K^{OL}_{α}) in the $\mathcal{S}_{train}$ set:

$$K^{OL}_{\alpha} = \frac{\left|(\mathbf{X},\mathbf{Y}) : (\mathbf{X}_i,\mathbf{Y}_i), i \in \mathcal{S}_{train}\right|}{|\mathcal{S}_{train}|} \times 100\%. \quad (4)$$

When conducting research, the values of this indicator varied from 0% (steganalytics do not have access to the encoder and can only use available stego images) to 100% (steganalytics can form stego images for arbitrary CI). This allows studying the influence of prior informationabout used embedding methods on the detection accuracy of SD [14].

The standard statistical model SPAM [15] was used to estimate parameters of analyzed DI, while the assignment of the analyzed image to classes of cover or stego images was carried out using random Forest classifier [8]. Also, the cases of usage cover rich model maxSRMd2 [6], as well as advanced artificial neural networks GB-Ras [4] and ASSAF [5] were considered. The maxSRMd2 model is based on usage HPF ensemble for DI pre-processing, and the further applying of Markov chain models to estimate correlation of adjacent pixels brightness [6]. The GB-Ras convolutional neural network allows ensuring high detection accuracy for wide range of SM due to usage of specialized convolutional methods in the input layers, namely depthwise separable convolution [4]. In contrast to the GB-Ras network, the hybrid ASSAF network is based on usage of DAE to estimate parameters of CI using the available (noisy) data, and further analysis differences between these parameters. This makes possible to ensure high detection accuracy for ASSAF network even under limited prior information about used SM [5].

4.2 Experiments

The study of the detection accuracy by variation of PPM types was carried out in several stages. At the first stage, the analysis of the achievable accuracy of the SD was done by using the considered pre-processing methods eqs. (2) and (3). The dependence of stego images detection error P_E on cover image payload for stego images formed according to the HUGO steganographic method, when using the methods eqs. (2) and (3) are presented in Fig. 2.

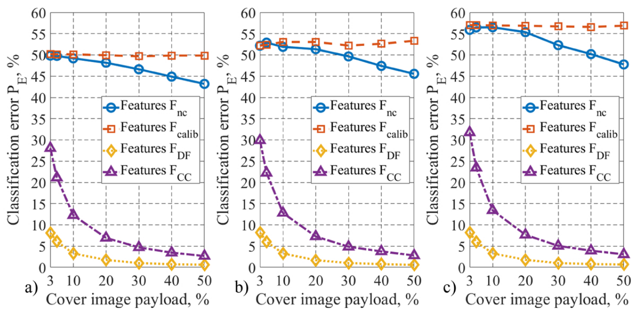

Fig. 2. Dependencies of stego images detection error P_E on cover image payload for stego images formed according to the HUGO steganographic method, when using preprocessing methods eqs. (2) and (3) for the ALASKA package and value of K^{OL}_{α} parameter: (a) — $K^{OL}_{\alpha} = 100\%$; (b) — $K^{OL}_{\alpha} \in \mathcal{U}(0;100)$; (c) — $K^{OL}_{\alpha} = 0\%$.

The decrease of the K^{OL}_{α} value from 100% to 0% leads to corresponding increase of P_E by 7% (Fig. 2). This is caused by a gradual reducing the ratio of cover-to-stego pairs in the $\mathcal{S}_{train}$ set.

Applying of $\mathbf{F}_{calib}$ vectors leads to slight changes of P_E values by varying of cover image payload (Fig. 2). This is caused by the fact that the statistical characteristics of the DI after applying of preprocessing methods eqs. (2) and (3) will coincide with the corresponding characteristics of the cover or stego images. As a result, the detection accuracy of SD will be determined only by the parameters of used feature extraction and classification methods.

On the other hand, applying of the considered methods eqs. (2) and (3) leads to the proportional to changes of elements both for both $\mathbf{F}_{DF}$ vectors and statistical parameters of CI caused by message hiding. As a result, the detection accuracy significantly increases (Fig. 2) especially for the case of low cover image payload (less than 10%), where the efficiency of modern SD is usually low.

The use of vectors $\mathbf{F}_{CC}$ during SD tuning also leads to decrease of P_E values compared to the case of using $\mathbf{F}_{nc}$ vectors (Fig. 2). This is explained by the fact that $\mathbf{F}_{CC}$ vectors include statistical characteristics of both initial and processed images, which leads to a doubling of the number of data elements of the vectors and corresponding increasing of complexity of SD adjustment.

For comparison, Fig. 3 shows the dependence of the values of the classification error P_E on cover image payload for stego images formed according to the MiPOD steganographic method, when using image preprocessing methods eqs. (2) and (3).

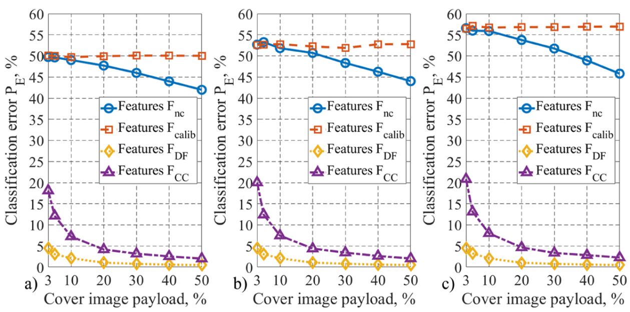

Fig. 3. Dependencies of the error values P_E on cover image payload for stego images formed according to the MiPOD steganographic method, when using preprocessing methods (2)-(3) for the ALASKA package and the value of the parameter K_α^{OL}: (a) — $K_\alpha^{OL} = 100\%$; (b) — $K_\alpha^{OL} \in \mathcal{U}(0;100)$; (c) — $K_\alpha^{OL} = 0\%$.

The obtained results for MiPOD embedding method (Fig. 3) confirm the previously obtained results for the HUGO method (Fig. 2), namely significant increase of detection accuracy by using image preprocessing methods eqs. (2) and (3). It is worth to note the reduction of P_E values on $5\% - 7\%$ by using $\mathbf{F}_{DF}$ and $\mathbf{F}_{CC}$ vectors for the MiPOD method compared to the corresponding values for the HUGO method (Fig. 2). This indicates about greater changes of adjacent pixels brightness correlation by data embedding, according to the novel steganographic method MiPOD, compared to the case of using the HUGO embedding method.

At the second stage of research, the comparative analysis of detection accuracy for stego images formed according to the considered embedding methods was performed. The case of using both preprocessing methods eqs. (2) and (3) and modern SD was considered. The value of the detection error P_E for stego images, formed according to the HUGO and MiPOD steganographic methods, for SD based on usage of pre-processing methods eqs. (2) and (3), as well as statistical models SPAM [15] and maxSRMd2 [6], artificial neural networks GB-Ras [4] and ASSAF [5] are given in table 1.

Table 1 The value of detection error P_E for stegodetectors based on usage of proposed image preprocessing methods $\mathsf{F}_{ideal}^{CE}(\mathbf{X},\mathbf{Y})$ and $\mathsf{F}_{ideal}^{SE}(\mathbf{X},\mathbf{Y})$, statistical model maxSRMd2 and artificial neural networks GB-RAS and ASSAF at $K_\alpha^{OL} = 0\%$ for the ALASKA image dataset

Type of stegdetector	Cover image payload, %			
	5%	10%	30%	50%
HUGO embedding method				
$\mathbf{F}_{DF}$	6.02	3.30	1.01	0.67
$\mathbf{F}_{CC}$	23.45	13.43	5.06	3.07
SPAM	56.46	56.50	52.29	47.76
maxSRMd2	47.33	44.44	35.04	28.59
GB-Ras	49.51	48.73	46.27	43.67
ASSAF	11.99	11.70	10.56	9.63
MiPOD embedding method				
$\mathbf{F}_{DF}$	3.25	2.07	0.79	0.49
$\mathbf{F}_{CC}$	13.11	8.01	3.37	2.28
SPAM	55.97	55.88	51.75	45.80
maxSRMd2	48.39	45.84	36.52	29.50
GB-Ras	49.32	48.57	46.16	44.20
ASSAF	12.18	12.09	12.03	12.14

The use of considered image pre-processing methods eqs. (2) and (3) makes it possible to reduce the value of detection error P_E up to 9 times in comparison with modern SD based on cover rich models and ANN. At the same time, the revealed decreasing of P_E values is also preserved even for low cover image payload (less than 10%), which is of particular interest when performing the stegoanalysis of advanced embedding methods.

It is worth to note that usage of special types of PPM, namely denoising autoencoder, allows approaching the achievable accuracy of stego images detection when using $\mathsf{F}_{ideal}^{CE}(\mathbf{X},\mathbf{Y})$ and $\mathsf{F}_{ideal}^{SE}(\mathbf{X},\mathbf{Y})$. This proves the prospects of using DAE for pre-processing of analyzed images to increase detection accuracy of SD.

Conclusion

The estimation of achievable detection accuracy of stego images by varying the type of image preprocessing methods and limitations of priori data about SM was performed. The optimal methods of preprocessing of DI based on the criterion of minimizing the stego images detection error are determined. It was revealed that applying of image-preprocessing methods aimed at restoration of cover image based on the available (noisy) data and the isolation of DI changes caused by stegodata embedding allows approaching to estimated achievable detection accuracy. It was shown that application of these methods allows reducing the detection error value for stego images formed according to modern HUGO and MiPOD steganographic methods up to 9 times, even for the case of low cover image payload (less than 10%) and the limitations of a priori data regarding the characteristics of embedding methods.

It has been established that the use of specialized image preprocessing methods, namely denoising autoencoders, allows ensuring detection accuracy, which is comparable to the use of the proposed preprocessing methods. Thus, research effectiveness of stegdetectors based on DAE for processing real DI, which are characterized by high variability of statistical parameters, is of further interest.

References

[1] Hassaballah M. (2020) *Digital Media Steganography: Principles, Algorithms, and Advances*, Academic Press, 386 p., ISBN: 978-0128194386.

[2] Goodfellow I., Bengio Y. and Courville A. (2016) *Deep Learning*, The MIT Press, 800 p., ISBN: 978-0262035613.

[3] Boroumand M., Chen M. and Fridrich J. (2019). Deep Residual Network for Steganalysis of Digital Images. *IEEE Transactions on Information Forensics and Security*, Vol 14, No. 5, p. 1181-1193, DOI: 10.1109/TIFS.2018.2871749.

[4] Tabares-Soto R., Arteaga-Arteaga H. B., Bravo-Ortiz M. A., Mora-Rubio A., Arias-Garzón D., Alzate-Grisales J. A., Burbano-Jacome A. B., Orozco-Arias S., Isaza G. and Ramos-Pollán R. (2021). GBRAS-Net: A Convolutional Neural Network Architecture for Spatial Image Steganalysis. *IEEE Access*, Vol 9, p. 14340-14350. DOI: 10.1109/ACCESS.2021.3052494.

[5] Cohen A., Cohen A. and Nissim N. (2020). ASSAF: Advanced and Slim StegAnalysis Detection Framework for JPEG images based on deep convolutional denoising autoencoder and Siamese networks. *Neural Networks*, Vol 131, p. 64-77. DOI: 10.1016/j.neunet.2020.07.022.

[6] Denemark T., Sedihi V., Holub V., Cogranne R. and Fridrich J. (2014). Selection-channel-aware rich model for Steganalysis of digital images. *International Workshop on Information Forensics and Security*, Atlanta, USA. DOI: 10.1109/WIFS.2014.7084302.

[7] Kodovsky J. and Fridrich J. (2009) Calibration revisited, *11th ACM workshop on Multimedia and security*, ACM, p. 63-74, DOI: 10.1145/1597817.1597830.

[8] Kodovsky J., Fridrich J. and Holub V. (2011) Ensemble Classifiers for Steganalysis of Digital Media. *IEEE Transactions on Information Forensics and Security*, Vol. 7, Issue 2, p. 432-444, DOI: 10.1109/TIFS.2011.2175919.

[9] Progonov D. (2021) Statistical stegdetectors performance by message re-embedding. *Theoretical and Applied Cybersecurity*, Vol.3, No. 1, p. 5-14. DOI: 10.20535/tacs.2664-29132021.1.251291.

[10] Progonov D.O. (2021) Effectiveness of stego images pre-noising with fractional noise for digital image steganalysis. *Applied Aspects of Information Technology*, Vol. 4, issue 3, p. 261-270, DOI: 10.15276/aait.03.2021.5.

[11] Cogranne R., Giboulot Q. and P. Bas. (2019) The ALASKA Steganalysis Challenge: A First Step Towards Steganalysis. *ACM Workshop on Information Hiding and Multimedia Security*, New York, USA, p. 125-137, DOI: 10.1145/3335203.3335726.

[12] Filler T. and Fridrich J. (2010) Gibbs Construction in Steganography. *IEEE Transactions on Information Forensics and Security*, Vol. 5, Issue 4, p. 705-720, DOI: 10.1109/TIFS.2010.2077629.

[13] Sedighi V., Cogranne R. and Fridrich J. (2015) Content-Adaptive Steganography by Minimizing Statistical Detectability. *IEEE Transactions on Information Forensics and Security*, Vol. 11, Issue 2, DOI: 10.1109/TIFS.2015.2486744.

[14] Progonov D. (2020) Performance of Statistical Stegdetectors in Case of Small Number of Stego Images in Training Set. *IEEE International Scientific-Practical Conference "Problems of Infocommunications Science and Technology"*, Kharkiv, Ukraine, DOI:10.1109/PICST51311.2020.9467901.

[15] Pevny T., Bas P. and Fridrich J. (2010) Steganalysis by Subtractive Pixel Adjacency Matrix. *IEEE Transactions on Information Forensics and Security*, Vol. 5, Issue 2, p. 215-224, DOI: 10.1109/TIFS.2010.2045842.

Дослідження впливу методів попередньої обробки цифрових зображень на точність виявлення стеганограм

Прогонов Д. О.

Особливістю сучасних методів виявлення несанкціонованої (прихованої) передачі конфіденційних даних в інформаційно-комунікаційних мережах є широке використання методів попередньої обробки досліджуваних файлів, зокрема цифрових зображень. Метою даних методів є детектування слабких змін статистичних параметрів зображення-контейнеру, обумовлених прихованням повідомлень. Значна кількість даних методів заснована на використанні ансамблів фільтрів високих частот, що дозволяє забезпечити високу точність виявлення стеганограм (більше 95%), сформованих згідно відомих стеганографічних методів. Проте вагомим обмеженням практичного застосування запропонованих методів попередньої обробки цифрових зображень є висока обчислювальна складність процедури формування даних ансамблів для мінімізації помилки виявлення стеганограм. Це унеможливлює швидке переналаштування стегодетекторів для виявлення стеганограм, сформованих згідно апріорно невідомих стеганографічних методів. Тому становить інтерес розробка швидких методів попередньої обробки досліджуваних зображень, здатних надійно виявляти слабкі зміни статистичних параметрів зображення-контейнеру в умовах обмеженості апріорних даних щодо використаного стеганографічного методу. Робота присвячена дослідженню досяжної точності роботи стегодетектору при варіації методів попередньої обробки цифрових зображень. За результатами дослідження визначено оптимальні методи попередньої обробки зображень для мінімізації помилки виявлення стеганограм. Дані методи дозволяють суттєво (до 9 разів) зменшити помилку класифікації стеганограм у порівнянні з сучасними методами попередньої обробки зображень, навіть у найбільш складному випадку слабкого заповнення зображення-контейнеру стегоданими (менше 10%) та обмеженості апріорних даних щодо використаного стеганографічного методу. Виявлено, що використання спеціалізованих методів обробки зображень, а саме знешумлюючих автоенкодерів, дозволяє наблизити точність роботи стегодетекторів до отриманих оцінок досяжної точності роботи стегодетекторів.

Ключові слова: стегоаналіз; методи попередньої обробки стеганограм; цифрові зображення