Joint multi-domain feature learning for image steganalysis based on CNN

Ze Wang
Guizhou University
Mingzhi Chen
Beijing Information Science and Technology University
Yu Yang (yangyu@bupt.edu.cn)
https://orcid.org/0000-0002-5045-804X
Min Lei
Beijing University of Posts and Telecommunications
Zhexuan Dong
University of California, Irvine, CA, U.S.

Research

Keywords: Image steganalysis, Convolutional neural networks, Feature learning, Joint domain,Nonlinear detection

Posted Date: March 15th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-17286/v1

Version of Record: A version of this preprint was published on July 8th, 2020. See the published version at https://doi.org/10.1186/s13640-020-00513-7.

RESEARCH

Joint multi-domain feature learning for image steganalysis based on CNN

Ze Wang[1,2], Mingzhi Chen[3], Yu Yang[2,1*], Min Lei[2,1] and Zhexuan Dong[4]

Abstract

Recently steganalysis methods based on convolutional neural networks (CNN) have achieved great improvement. However, detection against adaptive steganographic algorithms with low embedding rates has still been a challenging task. To deal with this problem, we propose a CNN steganalysis model employing the joint domain detection mechanism and nonlinear detection mechanism. For the joint domain detection mechanism, we use not only the high-pass filters from the SRM for spatial residuals, but also the patterns from the DCTR for frequency steganographic impacts. For the nonlinear detection mechanism, we enlarge steganographic effects by nonlinearly transforming the extracted steganographic residual information. In addition, we innovatively put forward a model learning method based on the high learning ability of a model. That is, we use lower embedding rate image datasets to train a model and higher embedding rate image datasets to test the model, which effectively improves sensitivity to steganographic traces. Compared with the existing steganalysis models such as SRM+EC, Ye-Net, Xu-Net, Yedroudj-Net and Zhu-Net, the detection accuracy of our model is about 4%∼6% higher than that of the best Zhu-Net model.

Keywords: Image steganalysis; Convolutional neural networks; Feature learning; Joint domain; Nonlinear detection

Introduction

Steganalysis [1] and information hiding are restricted and promoted mutually [2], [3]. And there is a more hopeful prospect to carry out the steganalysis work.

Image steganography is a technique that hides secret messages in images. In the frequency domain, the image is converted to frequency domain by discrete cosine transform (DCT) [4], discrete wavelet transform

*Correspondence: yangyu@bupt.edu.cn

[2]Laboratory of Cyberspace Security, Beijing University of Posts and Telecommunications, 100876 Beijing, China

[1]State Key Laboratory of Public Big Data, Guizhou University, 550025 Guizhou Guiyang, China

Full list of author information is available at the end of the article

(DWT) [5] and so on. Then the secret messages are embedded in the transformation coefficients. The popular algorithms are J-UNIWARD [6], nsF5 [7], UED [8], UERD [9]. In the spatial domain, the steganography algorithm is characterized by directly changing the pixels. The typical algorithms are the least significant bit (LSB) [10, 11], LSB matching [12], pixel-value differencing (PVD) [13]. The above algorithms can be regarded as the non-adaptive steganography algorithms. The adaptive steganography algorithms has been proved to have better performance. Currently popular adaptive algorithms are edge adaptive image steganography (EA) [14], HUGO [15], HILL [16], MiPOD [17], S-UNIWARD [6]. Now the security of steganography algorithms keeps on increasing [18], [19], [20], so the detection encounters more challenges.

The traditional steganographic analysis is mainly based on the manual feature extraction including gray level co-occurence matrix (GLCM) [21], Local Binary Patterns [22] and so on. The Spatial Rich Model (SRM) [23] is one of the most successful spatial steganographic detection algorithm. For a long time, scholars propose lots of steganographic analysis algorithms with SRM's ideas, so there is no major steganographic analysis achievements.

With the development of Graphics Processing Unit (GPU) parallel computing and deep learning, scholars hold a view that learning characteristics from the data itself can be more effective and gradually achieve it. In 2014, Tan and Li [24] proposed the first steganalysis model that applied deep learning techniques. In 2015, Qian et al. [25] proposed the first convolutional neural network model using supervised learning methods. In 2016, Xu et al. [26] proposed a CNN model similar to Qian's model. The difference was that they used an absolute value layer (ABS) and a 1x1 convolution kernel. At this time, Qian et al. [27] proposed to use the idea of transfer learning to improve the feature learning. The above models were all implemented in the spatial domain. In 2017, there were many steganographic analysis achievements in the frequency domain. Zeng et al. [28, 29] proposed a JPEG-based steganographic analysis model. Xu et al. [30], inspired by ResNet, proposed a new CNN model consisting of 20 convolutional layers with BN. Ye et al. [31] proposed a spatial domain CNN model. And they added a truncated linear unit (TLU) activation function to the preprocessing layer. The main trend in 2017 was to optimize the convolutional neural network architecture through ResNet [32], and draw on the feature extraction method of SRM. In 2018, Yedroudj et al. [33] proposed a spatial domain CNN model consisting of five convolutional layers. In addition to the traditional image datasets BOSSBass [34], they added the BOWS2 [35] image datasets. Tsang et al. [36] improved Ye's network model, which made the model perform steganalysis on high-resolution images. Zhang et al. [37] proposed a new CNN model, and they used the depth separable convolution network and spatial pyramid pooling (SPP) to obtain the channel correlation and adapt to different sizes of images.

The nonlinear characteristic of SRM has the perfect theoretical research achievement, but it has not been applied. The research shows that the combination of spatial and frequency domain information can be more conducive to detect spatial steganographic feature, but this idea has not been applied. Previous studies have shown that different embedding rate steganog-

raphy operations have similar influence modes of image neighboring pixel correlation, but this idea has not been used by most people.

In this paper, we propose a novel spatial domain steganalysis model called Wang-Net. It has the following characteristics:

(1) We firstly apply the joint domain detection mechanism, which can effectively improve the detection performance.

(2) We firstly apply the nonlinear feature detection mechanism, and extracts the nonlinear features by performing MinMax nonlinear operations on the residual information.

(3) We propose and apply a novel model transfer learning method, that is, we use low embedding rate images to train high learning ability model, and use high embedding rate images to verify and test.

The rest of this article is organized as follows. In Section 2, we introduce the SRM in the spatial domain and the DCTR in the frequency domain. In Section 3, we propose our CNN model. Our simulational results and analysis are in Section 4. Finally, we made a summary in Section 5.

Preliminaries

We mainly extract feature information through HPFs from SRM and DCT patterns from DCTR. The contents of SRM and DCTR are as follows.

SRM

We can see the feature extraction method of SRM steganalysis in Fig. 1. Firstly, the residual map submodels are obtained by the high-pass filter, then the fourth-order co-occurrence matrix of each residual map submodel is extracted by quantization, rounding and truncation. Finally, the elements of these co-occurrence matrices are rearranged to form the steganalysis feature vector.

Scholars have designed a variety of HPFs in SRM, and they use them to generate the residual map submodel, which is as follows:

$$R_{mn} = pred(N_{mn}) - cI_{mn} \tag{1}$$

Where c is called the residual order, m and n represent the pixel coordinates, N_{mn} is the adjacent pixel of I_{mn}, and $pred(N_{mn})$ is the predictor of cI_{mn}. Generally, the number of pixels in N_{mn} is equal to c.

The residuals mainly include first-order, second-order, third-order, SQUARE, EDGE3x3 and EDGE5x5 six types, and each type of residuals is divided into linear filtering residuals and nonlinear filtering residuals. The typical residuals and high-pass filters can be seen in Table 1, Eq.(2), Eq.(3), Eq.(4). The residuals in the horizontal, vertical, diagonal, and anti-angular directions are denoted as R^h,R^v,R^d,R^m.

$$\begin{aligned} R_{ij}^{min} &= min(R_{ij}^h, R_{ij}^v, R_{ij}^d, R_{ij}^m), \\ R_{ij}^{max} &= max(R_{ij}^h, R_{ij}^v, R_{ij}^d, R_{ij}^m) \end{aligned} \tag{2}$$

$$\begin{bmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 2 & -2 & 2 & -1 \\ 2 & -6 & 8 & -6 & 2 \\ -2 & 8 & -12 & 8 & -2 \\ 2 & -6 & 8 & -6 & 2 \\ -1 & 2 & -2 & 2 & -1 \end{bmatrix} \tag{3}$$

$$\begin{bmatrix} 2 & -1 \\ -4 & 2 \\ 2 & -1 \end{bmatrix}, \begin{bmatrix} -2 & 2 & -1 \\ 8 & -6 & 2 \\ -12 & 8 & -2 \\ 8 & -6 & 2 \\ -2 & 2 & -1 \end{bmatrix} \quad (4)$$

For the linear residual, Table 1 has given the calculation method of the first-order, second-order, third-order linear residuals. It is not difficult to find that the linear residuals of SQUARE, EDGE3x3 and EDGE5x5 only use more directional neighborhood pixels in the calculation. The SQUARE, EDGE3x3 and EDGE5x5 high-pass filters are shown in Eq.(3) and Eq.(4). In fact, the linear residual calculation method can be converted to the convolution operation:

$$R = I * K = (R_{ij}) = (\sum_{r,c} x_{i,j}^{r,c} k^{r,c}) \quad (5)$$

Where i, j are the pixel coordinates, and r, c are the index of the kernel.

As shown in Eq.(2), the nonlinear residual can be obtained by finding the maximum or minimum of some linear filtering residuals. As shown in Table 1, we take the first-order linear residual as the residual prototype, there are totally eight first-order linear residuals:

$$\begin{aligned} R_{ij} =& \{y_{i,j+1} - y_{ij}, y_{i+1,j+1} - y_{ij}, y_{i+1,j-1} - y_{ij}, \\ & y_{i,j+1} - y_{ij}, y_{i,j-1} - y_{ij}, y_{i-1,j+1} - y_{ij}, \\ & y_{i-1,j} - y_{ij}, y_{i-1,j-1} - y_{ij}\} \end{aligned} \quad (6)$$

Then the first-order nonlinear residual is:

$$R_{ij}^{min} = min\{R_{ij}\} \quad (7)$$

$$R_{ij}^{max} = max\{R_{ij}\} \quad (8)$$

The nonlinear residuals combine the statistical characteristics of the same kind of linear residuals, which fully reflect the adjacent pixels correlation changes in image caused by the steganography, and have extremely high application value.

DCTR

In the frequency domain, the discrete cosine transform Residual (DCTR) [38] is a general steganographic detection algorithm. The steps of its feature processing are as follows:

(1) Obtain 64 8x8 DCT bases patterns by calculation, then obtain feature maps by convoluting the decompressed JPEG image with the DCT basis patterns.

(2) Obtain the sub-feature maps by quantifying and truncating the raw feature maps.

(3) Compress the sub-feature maps as a 8000 dimensional feature vector.

In the above steps, the DCT basic patterns are 8x8 matrices,$B^{(i,j)} = (B_{mn}^{(i,j)})$, $0 \leq m,n \leq 7$, and $B_{mn}^{(i,j)}$ is calculated as follows:

$$B_{mn}^{(i,j)} = \frac{u_i u_j}{4} cos\frac{\pi i(2m+1)}{16} cos\frac{\pi j(2n+1)}{16} \quad (9)$$

Where $u_0 = \frac{1}{\sqrt{2}}$,$u_k = 1$ for $k > 0$.

DCT is defined as the convolution operation of the image and 64 DCT basic patterns $B^{(i,j)}$. In order to better understand DCT, we set the length and width of all images to a multiple of 8. Given a grayscale image $I \in R^{M \times N}$ of size $M \times N$ (M, N is a multiple of 8):

$$U(I) = \{U^{(i,j)} | 0 \leq i, j \leq 7\} \quad (10)$$

$$U^{i,j} = I * B^{i,j} \tag{11}$$

Where $U^{(i,j)} \in R^{(M-7)\times(N-7)}$, $*$ denotes a non-padded convolution operation.

Method

As shown in Fig. 2, our network model applies the joint domain and nonlinear detection mechanism. In the process of feature extraction, we simulate the SRM feature extraction method in the spatial domain and the DCTR feature extraction method in the frequency domain, and increase the extraction method of nonlinear residual features. we design nine different convolution layers. We also design the fully connected layer as the tenth layer for steganographic detection.

Nonlinear detection and joint domain detection mechanism in preprocessing layer

We apply the joint domain detection mechanism, that is, now our feature extraction method combines the advantages of spatial and frequency domain feature extraction methods. And the image spatial residual feature is obtained by SRM HPFs, the image frequency domain feature information is obtained by DCT patterns. At the same time, based on the more comprehensive characteristics of nonlinear residual statistical features, we apply the nonlinear detection mechanism.

First, we compare the proposed nonlinear detection mechanism with the commonly used spatial linear detection method. We use six types of SRM HPFs including the first-order, second-order, third-order, SQUARE, EDGE3x3, and EDGE5x5, whose amounts of high-pass filters are 8, 4, 8, 2, 4, and 4, respectively. And we obtain 30 linear residual feature maps through 30 high-pass filters. SQUARE is divided into SQUARE3x3 and SQUARE5x5, and SQUARE3x3 and EDGE3x3 belong to the same category in essence. So there are two nonlinear residual feature maps in SQUARE3x3 and EDGE3x3. There are also two nonlinear residual feature maps in SQUARE5x5 and EDGE5x5. And we obtain a total of 10 nonlinear residual feature maps by statistics. We did the steganalysis simulations on the single linear residual feature net (Linear Kernel-Net) and nonlinear residual feature net (Non-linear Kernel-Net).

According to Table 2, we know that the accuracy of Non-linear Kernel-Net is about 2%∼6% lower than that of Linear Kernel-Net in the simulations. We believe that our nonlinear residual features roughly calculate the maximum and minimum values of each type of linear residual features, and do not consider the distribution characteristics of residual feature values. Therefore, the adjacent pixel correlation changes in the image caused by steganography are not comprehensively reflected, that is, the advantage of the nonlinear residual feature is not fully utilized. However, the accuracy of Non-linear Kernel-Net is higher than that of CNN Steganography Analysis Network Ye-Net, indicating that Non-linear Kernel-Net still has good competition and can enhance the feature representation. Therefore, we add linear and nonlinear residual features to our network named All Kernel-Net.

According to the information in Table 3, All Kernel-Net has better steganographic analysis performance than Linear Kernel-Net and Non-linear Kernel-Net. In the WOW and S-UNIWARD steganography algorithms with the embedding rate of 0.2, the steganographic analysis accuracy of All Kernel-Net is about

0.3%∼%6 higher than that of Linear Kernel-Net and Non-linear Kernel-Net, and All Kernel-Net has no significant effect. Compared with the high embedding rate steganographic analysis accuracy, we know that this situation is mainly caused by too little embedded secret information. And All Kernel-Net combine the advantages of linear and nonlinear residual features, and has great steganalysis effect.

Based on All Kernel-Net, we apply the idea of joint domain detection and simulate the feature extraction method DCTR in the frequency domain, and we add it to the network.

According to the domain knowledge, small convolution kernels can effectively reduce the parameters scale, besides, matrix operations can take full advantage of the parallel computing of GPU. Therefore, we use a matrix with the channel number of 94 and the size of 5x5 instead of 94 convolution kernels, and initialize the matrix directly with DCT Patterns and HPFs. At this point, the calculation formula for the new DCT patterns is modified as follows:

$$B_{mn}^{(i,j)} = \frac{u_i u_j}{5} cos\frac{\pi i(2m+1)}{10} cos\frac{\pi j(2n+1)}{10} \quad (12)$$

Where $u_0 = 1, u_k = \sqrt{2}$ for $k > 0$, $0 \leq m, n \leq 4$, $0 \leq i, j \leq 7$.

In this way, we add the matrix initialized by DCT Patterns and HPFs to the preprocessing layer. This network is called Wang-Net that combines the advantages of linear and nonlinear feature extraction in the spatial and frequency domains. We did the steganalysis simulation. According to the results, after combining the advantages of frequency domain feature extraction, Wang-Net's steganographic analysis accuracy for WOW and S-UNIWARD is about 2%∼3% higher than All Kernel-Net. It strongly shows that the steganography effect of Wang-Net has been greatly improved.

Therefore, our network model finally applies the joint domain and nonlinear detection mechanism to simulate the extraction of spatial domain, frequency domain, linear and nonlinear steganographic features.

Detailed design in network architecture

Our network receives an image of size 256x256 and outputs two types of labels. The CNN has a number of layers, including a pre-processing layer, eight convolutional layers for feature extraction, and a fully connected layer for result classification. In the preprocessing layer, we use a convolution kernel with the channel number of 94 and the size of 5x5, which is initialized by the SRM filters and DCT Patterns.

In the feature extraction process, 3x3 convolution kernels are used in layers 2, 3, 4, 8, and 9, and 5x5 convolution kernels are used in layers 5, 6, and 7. In each convolutional layer, we add BN, ReLU and TLU nonlinear activation functions. And we also add average pooling to the 4, 5 and 6 convolutional layers.

Experiments

The environments

In the simulation, we applied two well-known content adaptive steganography algorithms to evaluate the performance of the CNN models, which are WOW and S-UNIWARD. We use the Matlab code of these steganographic algorithms to implement image steganography. As proposed in other steganalysis papers, we use a randomly embedded key when applying the steganography algorithm, which is also in line with the actual steganography situation.

Datasets

The datasets used in the simulations is the BOSSBase 1.01. BOSSBase 1.01 contains 10,000 512x512 grayscale cover images, which have different texture features and are widely used in steganalysis. Due to the limitations of GPU computing resources, in the simulation, we scale the images of BOSSBase 1.01 to 256x256 (using 'imresize()' in matlab, the function parameter remains the default configuration).

In the simulation, we use the steganography algorithm and cover images to generate 10,000 corresponding stego images. We find that the model is too complex and has strong learning ability. In order to prevent overfitting, we need to allocate as lots of datasets as possible for training. Therefore, the data ratio of training datasets, verification datasets and test datasets is 8:1:1.

Implementation details

We apply the AdaDelta [39] to train the network model, which accelerates the convergence of the model. Due to GPU memory limitations, we set the mini-batch size to 16. We use an exponential decay method with a decay rate of 0.95, a decay step of 2000, and an initial learning rate of 0.4. We also use Xavier [40] to initialize the weights and biases in all convolution layers.

Results and discussions

In this section, we compare the performance of Wang-Net and existing spatial domain steganalysis models, such as SRM+EC, Xu-Net, Ye-Net, Yedroudj-Net, and Zhu-Net. Then we use the idea of transfer learning to give full play to the model's migration ability.

We compare the detection performance of Wang-Net and other steganalysis algorithms. The results are shown in the Table 4.

According to the results, Wang-Net performance is better than SRM+EC, Xu-Net, Ye-Net, and Yedroudj-Net, but the detection accuracy is about 2%∼3% lower than that of Zhu-Net.

Our model has the strong learning ability. Inspired by the idea of transfer learning, we propose a novel migration learning method, that is, we use the datasets with lower embedding rates for training and use the datasets with higher embedding rates for verification and testing. We compare the performance of the networks again.

According to the results in Table 5, after using our new training mode, the performance of Wang-Net has been greatly improved and has surpassed that of Zhu-Net. Now our CNN model has the best steganalysis detection performance.

CONCLUSION

In the field of steganalysis, it is of great significance for applying the CNN. In this paper, we propose a CNN steganalysis model with three great advantages. In the first place, we pioneer the joint domain detection mechanism, and simulate the manual feature extraction method in the spatial and frequency domain to solve the spatial steganographic analysis problem. Furthermore, we creatively apply the nonlinear feature detection mechanism to the network. The nonlinear features are obtained by nonlinear transformation of the linear residual features, which synthesizes the characteristics and more fully reflects the change of the adjacent pixel correlation caused by the steganog-

raphy. What's more, we apply a novel model migration learning method. According to the characteristics of high learning ability of Wang-Net, we propose a model migration learning method, which uses low embedding rate images for training and high embedding rate images for testing. The learning method makes our model fully adapt to the distribution characteristics of embedded information. A large number of simulations have proved that our model has the best steganographic analysis detection performance. However, the nonlinear feature extraction method of our model is too rough. And our model is not suitable for steganalysis of color image. These issues will be addressed in further research.

Declarations

Availability of data and material

The datasets generated or analysed during the current study are available in the [BOSSBase] repository, [http://agents.fel.cvut.cz/boss/index.php?mode=VIEW &tmpl=materials]

Competing interests

The authors declare that they have no competing interests.

Funding

Supported by the National Key R&D Program of China(2017YFB0802703), Major Scientific and Technological Special Project of Guizhou Province (20183001), Open Foundation of Guizhou Provincial Key Laboratory of Public Big Data (2018BDKFJJ014), Open Foundation of Guizhou Provincial Key Laboratory of Public Big Data (2018BDKFJJ019) and Open Foundation of Guizhou Provincial Key Laboratory of Public Big Data (2018BDKFJJ022).

Author's contributions

Yu Yang and Ze Wang designed the algorithm, Ze Wang carried out the experiments and wrote the manuscript. Yu Yang, Mingzhi Chen, Min Lei gave the suggestions on the structure of manuscript and participated in modifying. Zhexuan Dong read and approved the final manuscript.

Acknowledgements

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

Authors' information

Not applicable.

Author details

[1]State Key Laboratory of Public Big Data, Guizhou University, 550025 Guizhou Guiyang, China. [2]Laboratory of Cyberspace Security, Beijing University of Posts and Telecommunications, 100876 Beijing, China. [3]College of New Media, Beijing Institute of Graphic Communication, 102600 Beijing, China. [4]Department of Computer Science, University of California, Irvine, 92697 CA, U.S..

References

1. Kang, Y., Liu, F., Yang, C., Luo, X., Zhang, T.: Color image steganalysis based on residuals of channel differences. Computers, Materials & Continua **59**, 315–329 (2019). doi:10.32604/cmc.2019.05242
2. Shi, L., Wang, Z., Qian, Z., Huang, N., Puteaux, P., Zhang, X.: Distortion function for emoji image steganography. Computers, Materials & Continua **58**, 943–953 (2019). doi:10.32604/cmc.2019.05768
3. Meng, R., Rice, S.G., Wang, J., Sun, X.: A fusion steganographic algorithm based on faster r-cnn. Computers, Materials and Continua **55**, 1–16 (2018). doi:10.3970/cmc.2018.055.001
4. Holub, V., Fridrich, J.: Low-complexity features for jpeg steganalysis using undecimated dct. IEEE Transactions on Information Forensics & Security **10**(2), 219–228

5. waveslim: Discrete wavelet transform (dwt). Encyclopedia of Multimedia, 188–188 (2006)
6. Holub, V., Fridrich, J., Denemark, T.: Universal distortion function for steganography in an arbitrary domain. Eurasip Journal on Information Security **1**(1), 1 (2014)
7. Fridrich, J.: Statistically undetectable jpeg steganography:dead ends challenges, and opportunities. (2007)
8. Guo, L., Ni, J., Shi, Y.Q.: Uniform embedding for efficient jpeg steganography. IEEE Transactions on Information Forensics & Security **9**(5), 814–825 (2014)
9. Guo, L., Ni, J., Su, W., Tang, C., Shi, Y.: Using statistical image model for jpeg steganography: Uniform embedding revisited. IEEE Transactions on Information Forensics & Security **10**(12), 2669–2680
10. Johnson, N.F., Jajodia, S.: Exploring steganography: Seeing the unseen **31**(2), 26–34 (2008)
11. Fridrich, J., Goljan, M., Rui, D.: Detecting lsb steganography in color, and gray-scale images. Multimedia IEEE **8**(4), 22–28 (2001)
12. Mielikainen, J.: Lsb matching revisited. IEEE Signal Processing Letters **13**(5), 285–287
13. Wu, D.C., Tsai, W.-H.: A steganographic method for images by pixel-value differencing. Pattern Recognition Letters **24**(9-10), 1613–1626
14. Luo, W., Huang, F., Huang, J.: Edge adaptive image steganography based on lsb matching revisited. IEEE Transactions on Information Forensics & Security **5**(2), 201–214 (2010)
15. Pevný, T., Filler, T., Bas, P.: Using high-dimensional image models to perform highly undetectable steganography **6387**, 161–177 (2010)
16. Li, B., Wang, M., Huang, J., Li, X.: A new cost function for spatial image steganography. 2014 IEEE International Conference on Image Processing, ICIP 2014, 4206–4210 (2015). doi:10.1109/ICIP.2014.7025854
17. Sedighi, V., Cogranne, R., Fridrich, J.: Content-adaptive steganography by minimizing statistical detectability. IEEE Transactions on Information Forensics & Security, 1–1
18. Qu, Z., Wu, S., Wang, M., Sun, L., Wang, X.: Effect of quantum noise on deterministic remote state preparation of an arbitrary two-particle state via various quantum entangled channels. Quantum Information Processing **16**(12), 306 (2017). doi:10.1007/s11128-017-1759-8
19. Qu, Z., Cheng, Z., Liu, W., Wang, X.: A novel quantum image steganography algorithm based on exploiting modification direction. Multimedia Tools and Applications **78** (2018). doi:10.1007/s11042-018-6476-5
20. Qu, Z., Li, Z., Xu, G., Wu, S., Wang, X.: Quantum image steganography protocol based on quantum image expansion and grover search algorithm. IEEE Access **7**, 50849–50857 (2019). doi:10.1109/ACCESS.2019.2909906
21. Haralick, R., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans Syst Man Cybern **SMC-3**, 610–621 (1973)
22. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis & Machine Intelligence **24**(7), 971–987
23. Fridrich, J., Kodovsky, J.: Rich models for steganalysis of digital images. IEEE Transactions on Information Forensics and Security **7**(3), 868–882 (2012). doi:10.1109/TIFS.2012.2190402
24. Tan, S., Li, B.: Stacked convolutional auto-encoders for steganalysis of digital images. In: Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific, pp. 1–4 (2014). doi:10.1109/APSIPA.2014.7041565
25. Qian, Y., Jing, D., Wei, W., Tan, T.: Deep learning for steganalysis via convolutional neural networks. Proceedings of SPIE - The International Society for Optical Engineering **9409**, 94090–9409010 (2015)
26. Xu, G., Wu, H.Z., Shi, Y.Q.: Structural design of convolutional neural networks for steganalysis. IEEE Signal Processing Letters **23**(5), 708–712 (2016)
27. Qian, Y., Jing, D., Wei, W., Tan, T.: Learning and transferring representations for image steganalysis using convolutional neural network. In: 2016 IEEE International Conference on Image Processing (ICIP) (2016)
28. Zeng, J., Tan, S., Li, B., Huang, J.: Large-scale jpeg steganalysis using hybrid deep-learning framework. IEEE Transactions on Information Forensics & Security **13**(5), 1200–1214 (2016)
29. Zeng, J., Tan, S., Li, B., Huang, J.: Pre-training via fitting deep neural network to rich-model features extraction procedure and its effect on deep learning for steganalysis. Electronic Imaging **2017**(7), 44–49
30. Xu, G.: Deep convolutional neural network to detect J-UNIWARD. CoRR **abs/1704.08378** (2017). 1704.08378
31. Ni, J., Ye, J., YI, Y.: Deep learning hierarchical representations for image steganalysis. IEEE Transactions on Information Forensics & Security, 1–1
32. Boroumand, M., Chen, M., Fridrich, J.: Deep residual network for steganalysis of digital images. IEEE Transactions on Information Forensics and Security **14**(5), 1181–1193 (2019). doi:10.1109/TIFS.2018.2871749
33. Yedroudj, M., Comby, F., Chaumont, M.: Yedrouj-net: An efficient cnn for spatial steganalysis
34. BossBase. http://agents.fel.cvut.cz/boss/index.php?mode=VIEW&tmpl=materials (2019)
35. Bows2. http://bows2.eclille.fr/index.php?mode=VIEW&tmpl=index1

(2019)

36. Tsang, C.F., Fridrich, J.: Steganalyzing images of arbitrary size with cnns. Electronic Imaging (2018)
37. Zhang, R., Zhu, F., Liu, J., Liu, G.: Efficient feature learning and multi-size image steganalysis based on cnn
38. Holub, V., Fridrich, J.: Low-complexity features for jpeg steganalysis using undecimated dct. IEEE Transactions on Information Forensics & Security **10**(2), 219–228
39. Zeiler, M.D., : Adadelta: An adaptive learning rate method. Computer Science (2012)
40. Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. Journal of Machine Learning Research - Proceedings Track **9**, 249–256 (2010)

Tables

Figures

Table 1 First-order, second-order, third-order linear residuals in the horizontal direction.

Residual Type	HPF	Linear Residual
First-order	(1, -1)	$R^h_{ij} = y_{i,j+1} - y_{ij}$
Second-order	(1, -2, 1)	$R^h_{ij} = y_{i,j-1} - 2y_{ij} + y_{i,j+1}$
Third-order	(1, -3, 3, -1)	$R^h_{ij} = y_{i,j-1} - 3y_{ij} + 3y_{i,j+1} - y_{i,j+2}$

Table 2 THE PERFORMANCE OF Linear Kernel-Net, Non-linear Kernel-Net AND Ye-Net ON RESAMPLED IMAGES. THE INVOLVED NETWORKS ARE TRAINED AND TESTED ON BOSSBase.

Algorithms	Linear Kernel-Net	Non-linear Kernel-Net	Ye-Net
WOW(0.2bpp)	0.710181	0.697077	0.669
WOW(0.4bpp)	0.818548	0.788306	0.768
S-UNIWARD(0.2bpp)	0.661290	0.608871	0.600
S-UNIWARD(0.4bpp)	0.765625	0.733871	0.688

Table 3 THE PERFORMANCE OF Linear Kernel-Net, Non-linear Kernel-Net AND All Kernel-Net ON RESAMPLED IMAGES. THE INVOLVED NETWORKS ARE TRAINED AND TESTED ON BOSSBase.

Algorithms	Linear Kernel-Net	Non-linear Kernel-Net	All Kernel-Net
WOW(0.2bpp)	0.710181	0.697077	0.713710
WOW(0.4bpp)	0.818548	0.788306	0.844254
S-UNIWARD(0.2bpp)	0.661290	0.608871	0.668851
S-UNIWARD(0.4bpp)	0.765625	0.733871	0.791835

Table 4 THE PERFORMANCE OF Wang-Net AND OTHER STEGANALYSIS MODELS ON RESAMPLED IMAGES. THE INVOLVED NETWORKS ARE TRAINED AND TESTED ON BOSSBase.

Algorithms	WOW (0.2bpp)	WOW (0.4bpp)	S-UNIWARD (0.2bpp)	S-UNIWARD (0.4bpp)
SRM+EC	0.635	0.745	0.634	0.753
Xu-Net	0.676	0.793	0.609	0.728
Ye-Net	0.669	0.768	0.600	0.688
Yedroudj-Net	0.722	0.859	0.633	0.772
Zhu-Net	0.766	0.882	0.719	0.847
Wang-Net	0.748992	0.859879	0.691028	0.818548

Table 5 THE PERFORMANCE OF Wang-Net AND Yedroudj-Net AND Zhu-Net ON RESAMPLED IMAGES. THE INVOLVED NETWORKS ARE TRAINED AND TESTED ON BOSSBase.

Algorithms	Payload	Yedroudj-Net	Zhu-Net	Wang-Net
WOW	0.2	0.722	0.766	0.811996
	0.3	-	-	0.887601
	0.4	0.859	0.882	0.919859
S-UNIWARD	0.2	0.633	0.719	0.777218
	0.3	-	-	0.815020
	0.4	0.772	0.847	0.888105

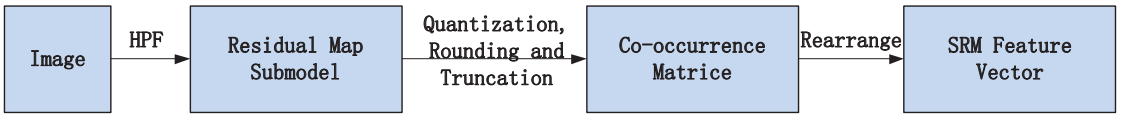

Figure 1 The feature extraction process of SRM steganalysis.

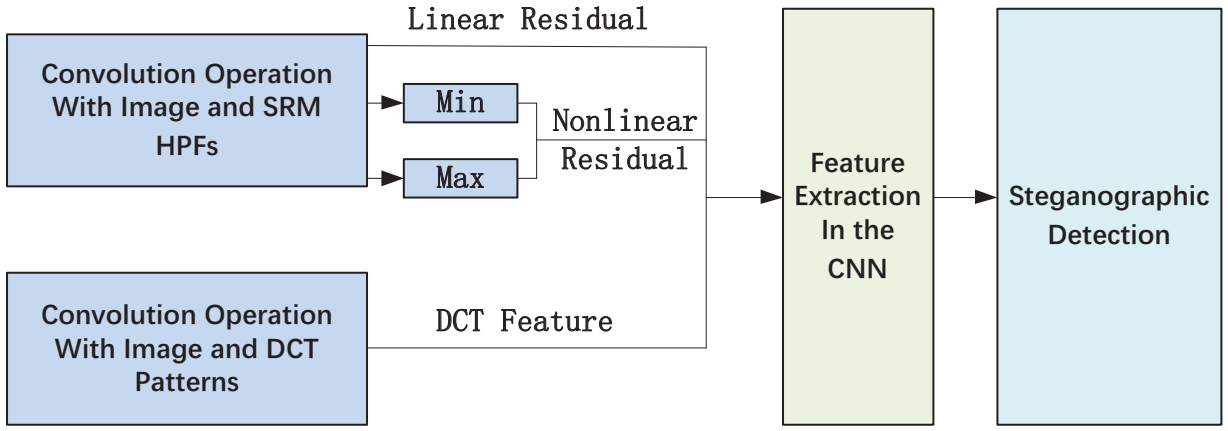

Figure 2 Integrated structure of the network.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- bmcarticle.bib
- spbasic.bst
- SRM1.eps
- 111.bib
- SRM1epsconvertedto.pdf
- bmcartbiblio.sty
- bmcart.cls
- 1bmcarticle.bbl
- kernelepsconvertedto.pdf
- 1bmcarticle.log
- bmcmathphys.bst
- 1bmcarticle.blg
- kernel.eps
- 1bmcarticle.tex
- 1bmcarticle.aux
- vancouver.bst
- 1bmcarticle.out