International Journal of Digital Crime and Forensics
Volume 13 • Issue 3 • Bi-Monthly 2021

A Modification-Free Steganography Algorithm Based on Image Classification and CNN

Jian Bin Wu, Central China Normal University, China
Yang Zhang, Central China Normal University, China
Chu Wei Luo, Central China Normal University, China
Lin Feng Yuan, Wuhan Maritime Communication Research Institute, China
Xiao Kang Shen, Central China Normal University, China

ABSTRACT

In order to improve the data-embedding capacity of modification-free steganography algorithm, scholars have done a lot of research work to meet practical demands. By researching the user's behavioral habits of several social platforms, a semi-structured modification-free steganography algorithm is introduced in the paper. By constructing the mapping relationship between small icons and binary numbers, the idea of image stitching is utilized, and small icons are stitched together according to the behavioral habits of people's social platforms to implement the graphical representation of secret messages. The convolutional neural network (CNN) has been used to train the small icon recognition and classification data set in the algorithm. In order to improve the robustness of the algorithm, the icons processed by various attack methods are introduced as interference samples in the training set. The experimental results show that the algorithm has good anti-attack ability, and the hiding capacity can be improved, which can be used in the covert communication.

KEYWORDS

Behavioral Habits, Convolutional Neural Network, Image Classification, Image Mosaic, Modification-Free Steganography

1. INTRODUCTION

Steganography is a technique of hiding the secret information in the carrier and extracting the secret information from the stego carrier, so as to achieve the purpose of the covert communication and copyright protection. Digital image is a common carrier which is often used in hiding information due to its large redundancy and wide application. For the traditional information hiding method, digital images are embedded with secret messages, which lead to the modification of the carrier itself. These modifications can cause some characteristics of the image to change. The third party determines whether the picture is embedded in the secret message by extracting these features. Though traditional information hiding has better robustness and larger capacity, it is difficult to resist steganography analysis and detection. In order to improve the security of covert communication, Modification-free Steganography algorithm has attracted extensive attention(Cao et al., 2018; Zhang et al., 2018; Zheng et al., 2017; Zhou et al., 2015). "Modification-free" steganography does not mean that no carrier is

DOI: 10.4018/IJDCF.20210501.oa4

needed, but it directly drives secret information to "Generate" or "Obtain" the stego carrier, The study of modification-free Steganography is divided into two directions: coding/mapping and deep learning.

Some scholars have studied coding/mapping modification-free steganography. It uses a certain feature of the image to establish a one-to-one mapping relationship with the binary sequence. Otrori and Kuriyama firstly proposed the idea of the data embedding in the texture synthesis process (Otori & Kuriyama, 2007; Otori & Kuriyama, 2009). Texture synthesis information hiding implements the information hiding in the process of texture synthesis, and the resulting large texture image is related to secret information. But the latest research shows that this method still has security holes(Zhou et al., 2016). Reference(Xu et al., 2015) proposes to use the geometric deformation to generate marbling effects. First, the secret information is directly written on the white paper, and then the background pattern and color coordinated with the color shape of the secret information are added to the blank portion, and finally the texture map is generated by using different deformation functions. But the information which is hidden in the above method is a text or a pattern with meaning, so it is not suitable for the hiding of binary data. But the above method has the problem of low practicality.

Other scholars have applied deep learning methods into information hiding. Volkhonskiy et al. first proposed the SGAN model, and used the anti-learning method to obtain the carrier image for steganography. The anti-learning is to make the cover image and the stego image closer to improve the steganography security. But the generated cover image is embedded using the traditional steganography method in the end(Volkhonskiy et al., 2017). The HayesGAN model proposed by Hayes uses the confrontation learning to directly generate the dense image. This method has a great improvement in security, but it cannot guarantee the complete extraction of the embedded secret information(Hayes & Danezis, 2017).

To address the above problems, this paper proposes a semi structured modification -free steganography algorithm based on the behavioral habits of social platform(Zhang et al., 2016). The specific implementation method is to splice small icons into pictures under the guidance of the text rules to achieve modification-free. Among them, the library is established on the basis of training, classification and recognition of small icons with the method of deep learning. CNN (Convolutional neural network) (Liu et al., 2018) is leveraged to extract the image semantics and to train it as a model input. The identification and classification of those small icons in the library are implemented in accordance with the high dimensional characteristic of images. Provided that the image may be attacked by the third party in the process of transmission, and the image data-set should contain all kinds of interfered samples for training. These samples are the images which are specially processed. The training set containing the interference samples can ensure that the trained CNN network can correctly classify the icon after the attack to the stego images, which strengthens the robustness of the algorithm.

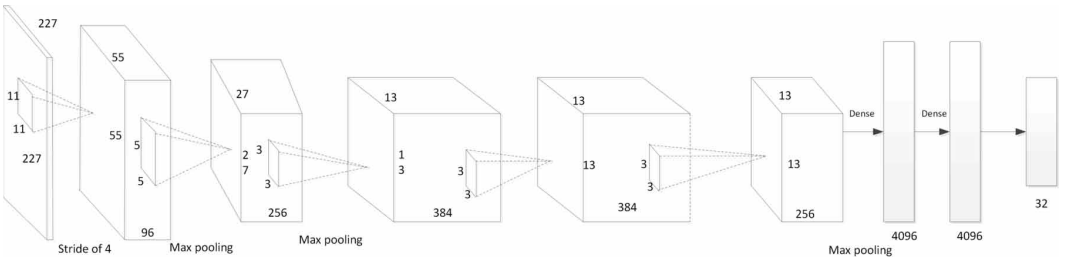
2. RELATED THEORIES

CNNs are widely used in deep learning for their low network model complexity and the ability to reduce the number of weights. Images are directly input into the network in the CNN based deep learning, which can avoid the process of data reconstruction and complex feature extraction in traditional algorithms, and improve the algorithm efficiency. Krizhevsky (2012) proposed the convolutional neural network Alex Net, which showed a better performance in the image classification and object detection than the traditional method. Although the performance of Google Net (Zegedy et al., 2015) and VGG (Simonyan, 2016) is better than that of the Alex Net network, the networks of the formers are much more complex, which takes much longer time for training. Considering the factors of training and recognition efficiency, this paper utilized Alex Net network for image recognition and classification. The network of which is relatively simple and easy to be trained. The performance can be improved by modifying some parameters in neural network.

We used Deep Convolutional Neural Network architecture, similarly to that (Krizhevsky et al., 2012), the first five of which are convolutional and other three are fully connected (see Figure 1). The

first, second and fifth convolutional layers are followed by max-pooling layers (Liu, 2018; Yong & Yao, n.d.). The first and second max-pooling layers are followed by the local response normalization layers. We used Rectified Linear Units (ReLU) as neurons (Hinton et al., n.d.; Sánchez, 2011). The first convolutional layer has 96 kernels of the size of 11*11*3 with a stride of 4 pixels. The second layer takes as input the max-pooled and response-normalized output of the first layer and filters it with 256 kernels of size 5*5*48. The third convolutional layer takes as input the max-pooled and response normalized output of the second layer and filters it with 384 kernels of size 3*3*256. The fourth layer has 384 kernels of size 3*3*192, and the fifth layer has 256 kernels of size 3*3*192. The fully-connected layers have 4096 neurons each. Max-pooling layers have the size of 3*3 and stride of 2. The final layer is 32-way Softmax.

Figure 1. The architecture of Net



3. PROPOSED METHOD

By considering the requirements from the capacity, bit error rate, robustness and rationality of the content of the stitched images, this paper aims to verify the feasibility of the algorithm. Therefore, the icon library only constructs 32 small icons; Meanwhile, the concept of the traditional modification-free steganography is leveraged, utilizing some features of the image to establish a mapping relationship with the binary sequence. This paper uses the semantics of the images to establish mapping relationships. The semantic label of each icon corresponds to a 5-bit binary sequence.

Since a stego image is made up of a series of small icons, it should conform to people's social habits in order to avoid being suspected by the third party. Therefore, the data mining method is used to mine and analyze various behavioral habits on the social platform; meanwhile, the information hiding is realized by using the stitching principle of the image to solve the problem. There are many types of rules for stitching icons, for example, adding an intellectual problem, which will make the stego image interesting. At present, five kinds of rules are designed to generate images with secret carrier, such as: "Calculating price", "Sub-category", "Making a story", "Finding the longest" and "Looking for the same." The stitched image has 5 templates of 2*2, 3*3, 4*4, 5*5and 6*6 in accordance with the length of the secret message. The final result of the stego image is shown in Figure 2.

3.1 The Embedding Algorithm

Figure 4 shows the block diagram of the embedding algorithm. The specific implementation steps are as follows:

Step 1: Segment the secret messages and use the length and serial number of each segment as the flag bit.

Figure 2. The stego image

Step 2: Add the flag bit to the beginning of each piece of the secret message, which is divided into small segments of 5-bit-length. Then select the certain icon in accordance with the mapping relationship between the 5-bit-sequence and the icon from the database.

Step 3: Select a stitching template based on the length of the secret messages and splice small icons in the order of the selected template. The 3*3 template is shown in the Figure 3.

Step 4: Choose a reasonable rule to describe the stitched image and add the rule below the stitched image.

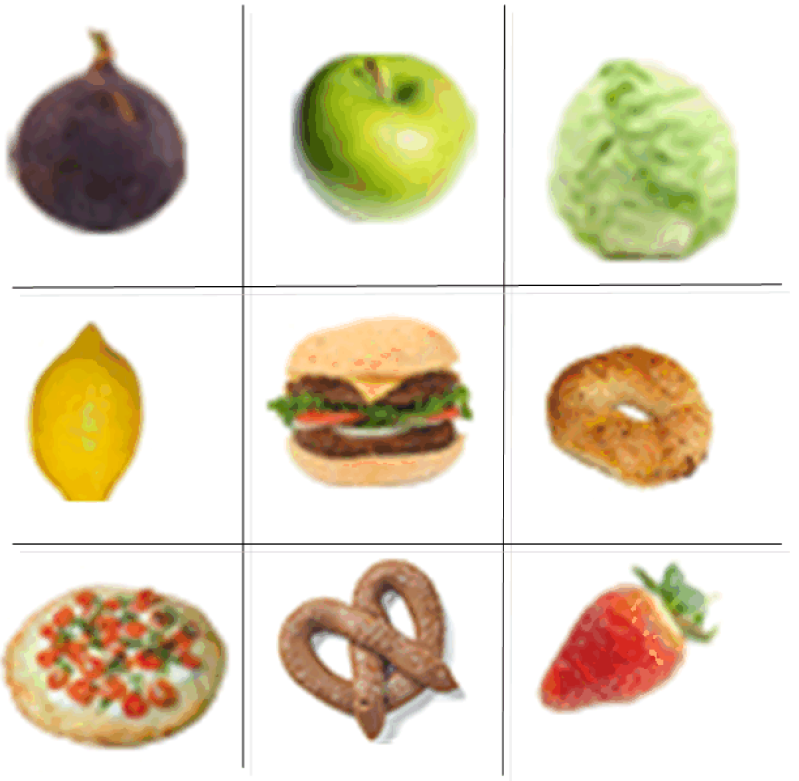
Step 5: Calculate a feature value of the image according to the information entropy and the weight of each sub-image of the picture:

$$E = \sum_{n=1}^{m} \frac{n^2}{n^2 - 1} M_n \tag{1}$$

Where M is the information entropy of each sub-image, and m is the number of small icons, and n is the sequence number of the small icon from top to bottom and from left to right.

Step 6: The feature values calculated in the previous step are embedded as a watermark into the expression of the rule by a DCT transform algorithm. Finally, the secret image is merged with

Figure 3. The 3*3 template



the small icon, and the feature watermark is embedded into the image which contains the rule to form a secret carrier(see Figure 5).

3.2 The Extracting Algorithm

Figure 6 shows the block diagram of the extracting algorithm. The following steps are executed to recover the original secret message:

Step 1: Detect the watermark. If the watermark exists, the secret messages will be extracted, otherwise it will not be proceeded to the next step.

Step 2: Split the secret image and preprocess the segmented image, then identify and classify the small icons, and obtain the binary sequence from the classification results according to the established mapping relationship.

Step 3: The flag bit in front of the binary sequence is extracted to obtain the length L of the secret message and the sequence number N of the small icons. The secret message is obtained according to L and N.

Figure 4. Secret information hiding flow diagram

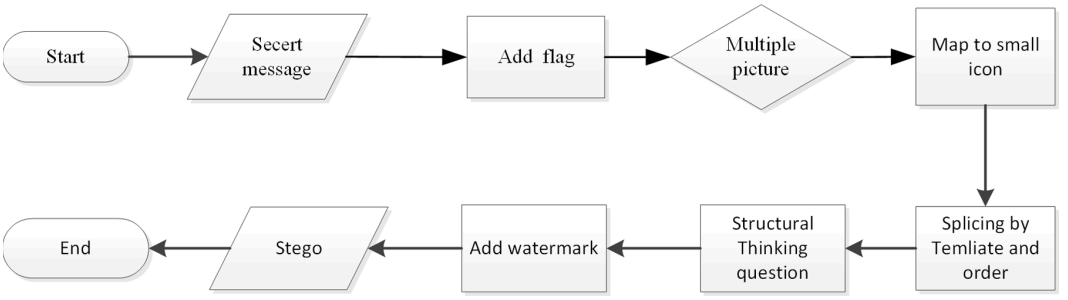
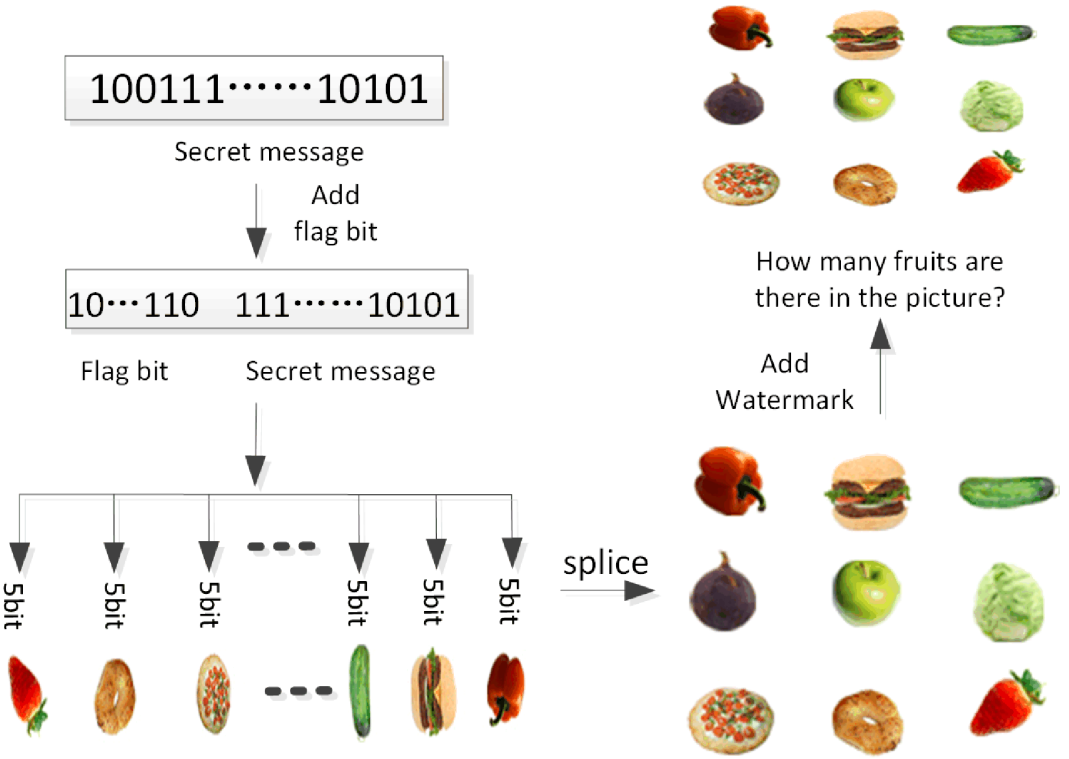


Figure 5. Schematic diagram of hiding secret information



4. ALGORITHM EXPERIMENT AND ANALYSIS

In order to verify the effectiveness of a semi-constructed modification-free algorithm proposed in this paper, the experiments are studied and analyzed in detail from three aspects: Robustness, security and algorithm capacity (Figure 7).

4.1 Robustness

This paper judges the robustness of the algorithm by the means of calculating its bit error rate. The bit error rate (BER) is defined as the ratio of the error number p of the decrypted information and the total number q of original secret information:

Figure 6. Secret information extract flow diagram

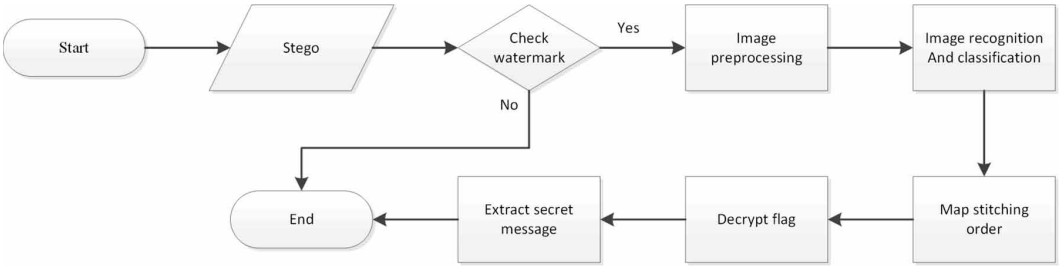


Figure 7. Schematic diagram of hiding secret information

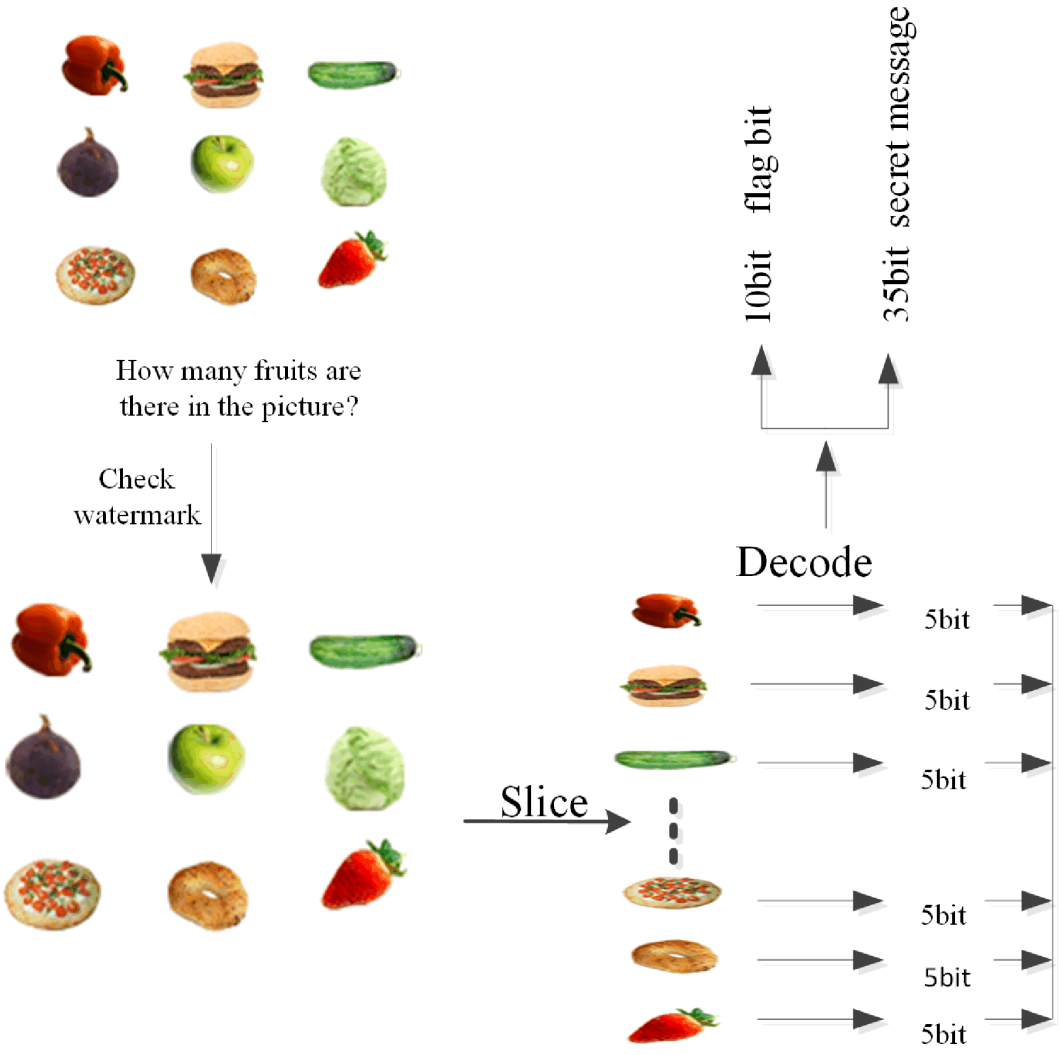


Table 1. BER comparison under salt and pepper noise attack

D	0.1	0.2	0.4	0.6	0.8	1.0
BER	0	0	0	0.15	0.42	0.9

$$BER = \mathrm{p} / q \times 100\% \tag{2}$$

In order to ensure that the small icons can be classified without any errors. In the training with Alex Net, the images were attacked by Gaussian noise, salt and pepper noise, mean filtering, JPEG compression. Therefore, When the salt & pepper noise intensity D<0.6, the bit error rate under the algorithm is 0. Once the noise intensity exceeds the threshold, the bit error rate will rush to a higher level (see Table 1).

The test results show that under the circumstances of changing compression quality and the intensity of Gaussian Noise, the BER of the algorithm in this paper is 0, which demonstrates that the algorithm is not interfered by compression attacks. so it is clear that the proposed algorithm has good robustness.

4.2 Security

In general, resistance to steganalysis is an important factor to judge whether an information hiding algorithm is of good quality. The success rate of the algorithm can be improved if and only if the anti-detection performance of which is improved. The traditional information hiding algorithm often leaves the modification traces when changing the carrier in the process of hiding the secret information, so it is difficult to resist the detection of various steganalysis tools. On this basis, this paper proposes a semi-structured modification-free steganography technology, which uses the mathematical-puzzle-like picture as the carrier to transmit the secret information. At the same time, the binary bit stream, which is transformed from the secret messages, is segmented into 5-bit-length sequence, and then these sequences are converted into small icons in accordance with the mapping relationship between the binary bit segment and the small icon. Finally, through deep learning and data mining, people's behavior habits on social platforms are extracted, which are the basis of that the construction rules and small icons are stitched according to the construction principles. In this way, the image representation of secret messages is completed. Since the stego image is not modified by the existing image, but is formed by splicing small icons. So the tradition steganographic analysis method for extracting image features cannot detect whether the stego image contains secret message. Because the image is spliced with certain rules and conforms to people's social habits, it will not cause third-part suspicion because of splicing. Therefore, the algorithm can resist the analysis and detection of various steganography tools, and it will not cause suspects from the third parties (Fridrich & Kodovsky, 2012; Goljan, 2001; Pevny & Fridrich, 2007).

4.3 Capacity

For an information hiding system, the hidden capacity is the maximum number of information bits that the selected carrier can hide in the process of transmission without being perceived. Since the secret message that can be expressed by a single picture is limited, the idea of splicing is used to splicing several pictures together to increase the capacity of each communication. The paper of the coverless information hiding algorithm which is based on the image coding and splicing proposed in the existing paper show that the number of bits that can be hidden in a single image which is 26(Wu et al., 2018), and A coverless information hiding Algorithm Based on Grayscale Gradient Co-occurrence Matrix proposes to use the gradient and gray level co-occurrence matrix to hide the secret messages. The number of the bits of the secret messages that can be hidden by a single picture is 8(Wu, 2018).

Table 2. Single picture hides the maximum number of bits

Method Hide the maximum number of bits
Paper (Wu, 2018) 8 Paper (Wu et al., 2018) 26 This paper 170

In the algorithm that is proposed in this paper, the more small icons contained in a stego image, the more secret messages that can be hidden. The picture showed in this article can contain up to 36 small icons, and each small icon can hide 5 bit. After removing the previous 10 bit flag, the number of the bits that can be hidden in a picture is 170. Table 2 shows a comparison among the reference (Wu et al., 2018), reference (Wu, 2018) and the proposed method.

From the experimental data in Table 1, it is evident that the proposed algorithm effectively improves the hidden capacity of the modification-free to a certain extent.

The experimental results show that the steganography algorithm of the paper has a certain degree of improvement in both capacity and bit error rate, and has high security.

5. CONCLUSION

In order to improve the communication efficiency and conceal capacity, this paper uses data mining technology to extract social behavior habits from social platforms, combining deep learning methods to train, identify and classify the icons. An icon library is built to construct the secret images with high behaviorally acceptance and the construction principles. A covert communication is implemented in an interesting and high-security way. The algorithm is tested in the MATLAB 2017b environment and Our network train on one Intel i5 6600 8GB CPUs. From the experimental results, the proposed algorithm has a good concealment and robustness, with high tolerance of attack from Gaussian noise, salt & pepper noise, JPEG compression, means filter and median filter. This method effectively improves the communication efficiency and capacity of secret messages. In our future work, a larger icon library will be built to expand the capacity of the algorithm, while behavioral habits mining is still a prevalent and roaring task.

ACKNOWLEDGMENT

This research was supported by Fund project: The National Nature Science Foundation of China U1736121.

REFERENCES

Cao, Y., Zhou, Z., Sun, X., & Gao, C. (2018). Coverless information hiding based on the molecular structu- re images of material. *Computers*, *Materials & Continua*, *54*(2), 197–207.

Fridrich & Kodovsky. (2012). Rich models for steganalysis of digital image. *IEEE Trans Information Forensics and Security, 7*(3), 868-882.

Goljan, M. (2001). *Lossless data embedding methods for digital images and detection of steganography*. State University of New York at Binghamton.

Hayes, J., & Danezis, G. (2017). Generating steganographic images via adversarial training. *Advances in Neural Information Processing Systems*, 1954–1963.

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. R. (n.d.). *Improving neural networks by preventing co-adaptation of feature detectors*. Academic Press.

Krizhevskya. (2012). *ImageNet classification with deep convolutional neural networks*. Advances in Neural Information Process in Systems.

Liu, J., Yuan, Q., & Yu, X. (2018). Review of convolutional neural networks. *Computer Era*, *11*, 19–23.

Liu, X. (2018). Transfer Learning Research and Algorithm Review. *Journal of Changsha University*, *32*(5), 29–36.

Otori, H., & Kuriyama, S. (2007). Data-embeddable texture synthesis. *Proceedings of the 8th Inter-national Symposium on Smart Graphics*, 146-157.

Otori, H., & Kuriyama, S. (2009). Texture synthesis for mobile data communications. *IEEE Computer Graphics and Applications*, *29*(6), 74–81. doi:10.1109/MCG.2009.127 PMID:24806781

Pevny, T., & Fridrich, J. (2007). Merging Markov and DCT features for multi-class JPEG steganalysis. *Proceedings of the Society for Photo-Instrumentation Engineers*, *6505*, 3–4. doi:10.1117/12.696774

Sánchez, J. F. (2011). High-dimensional signature compression for large-scale image classification. *Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*, 1665–1672.

Simonyan, K. (2016). *Very deep convolutional networks for Large-scale image recognition*. Available from: https://arxiv.org/abs/1409.1556

Xu, J., Mao, X., Jin, X., Jaffer, A., Lu, S., Li, L., & Toyoura, M. (2015). Hidden message in a deformation based texture. *The Visual Computer*, *31*(12), 1653–1669. doi:10.1007/s00371-014-1045-z

Volkhonskiy, D., Nazarov, I., & Borisenko, B. (2017). *Steganographic generative adversarial networks*. arXiv preprint arXiv:1703.05502.

Wu, Jia, & Liu. (2018). A Coverless Information Hiding Algorithm based on image coding and stitching. *Proceedings of 14th China Information Hiding Workshop*.

Wu, J. (2018). A Coverless Information Hiding Algorithm Based on Grayscale Gradient Co-occurrence Matrix. *IETE Technical Review*, 1–11.

Yong & Yao. (n.d.). Ren Biome, Search On Pooling Method of Convolution Neural Network. *Computer Engineering*. 10.19678/j.issn.10003428.0050129,2018-03-29

Zegedy, C., Liu, W., Jia, Y. Q., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE.

Zhang, X., Peng, F., & Long, M. (2018). Robust Coverless Image Steganography based on DCT and LDA Topic Classification. *IEEE Transactions on Multimedia*.

Zhang, X. P., Qian, Z. X., & Li, S. (2016). Prospect of Digital Steganography Research. *Journal of applied sciences*, *34*(5), 476–478.

Zheng, S., Wang, L., Ling, B., & Hu, D. (2017). Coverless Information Hiding Based on Robust Image Hashing. In Intelligent Computing Methodologies. Springer.

Zhou, H., Chen, K., Zhang, W., & Yu, N. (2016). Comments on steganography using reversible texture synthesis. *IEEE Transactions on Image Processing*. PMID:28252386

Zhou, Z., Sun, H., & Harit, R. (2015). Coverless Image Steganograph Without Embedding. In *International Conference on Cloud Computing and Security*. Springer.

Jianbin Wu is a Professor at the College of Physics Science and Technology, Central China Normal University. He received the MSc in Electronic Circuit and System and PhD in Communication and Information System from Wuhan University, in 2003 and 2009, respectively. His research interests include information security, information hiding, radar signal processing and low observable technology.

Yang Zhang received his BSc degree in Electronic and Information Engineering from Hubei University of Technology in 2016. He is currently a Master's candidate in Information Security at Central China Normal University, Wuhan, China. His research interests include information hiding and image processing.

Chuwei Luo received his BSc degree in Electronic and Information Engineering from Central China Normal University in 2017. He is currently a Master's candidate in Information Security at Central China Normal University, Wuhan, China. His research interests include information hiding and image processing

Linfeng Yuan works at Wuhan Maritime Communication Research Institute. His research interests include information security, information hiding.

Xiao Kang Shen is currently a Master's candidate in Information Security at Central China Normal University, Wuhan, China.