High Capacity Blind Information Hiding Based on Autoencoder

Zhenjie Bao*, Yadong Jin, Yue Liu, and Jingyun Hu

School of Information Engineering, Xizang Minzu University, Xianyang, Shanxi, China

*Corresponding author: zhenjibao@xzmu.edu.cn

Abstract. To realize image information hiding, we train a multi-layers autoencoder includes an encoder and a decoder. Using encoder architecture, we input the secret color image and generate the cover image where the size stays the same directly. In addition, we reconstruct the secret color image by inputting the stego images passing on the decoder which accomplish the secret image blind extraction. Our method realizes a high resolution and high payload color image steganography. Experimental results show that we can realize the high capacity blind specified secret images information hiding automatically by our method.

Keywords: Image information hiding, autoencoder, blind extraction, high capacity.

1. Introduction
With the development of Internet. The communication in the Internet is becoming more and more frequently. However, sometimes we want to transmit some secret images between the sender and the receiver through a public network channel which is full of sniffing and automatic detection. How can we transmit it like an innocent normal image and avoid to be noticed. We don’t want to arouse too much awareness from other people or automatic detection programs so that the communication can be in a safe and secret environment. Image information hiding is to hide secret images in an undetectable way. In traditional steganography algorithms, secret image was embedded into cover image. There were lots of stego-embedding algorithms, for example, the Least Significant Bit (LSB)\cite{1} matching in the spatial domain. Concealing secret messages based on image transformation such as DCT (Discrete Cosine Transform)\cite{2} and DWT (Discrete Wavelet Transform)\cite{3} had achieved impressive results. Besides, considering the content of image itself, adaptive steganography\cite{4} selected the embedding location of the cover image so that the security and visual quality of stego images were enhanced. The improvement in information hiding was various. Based on the molecular key, \cite{5} utilized the Ugi four-component reaction of perfluorinated acids to establish datasets and the molecular keys could be concealed by adsorption onto coffee and other things. Based on signal-pixel imaging (SPI), \cite{6} proposed a compressive optical steganography which sparse-sampled the secret image and then encoded it into the illumination patterns. By implementing an inverse Fourier transform on the result reflected light signals, the potential eavesdroppers could be reconstructed. Based on the chaotic random number generator, \cite{7} hid an image data with different color into color image in bit space. Based on modified histogram shifting, \cite{8} incorporated multi-bit secret data hiding. In recently
published literature, there are many new ideas in information hiding. By making use of the similarity of the pixels, [9] proposed a new optimization-based steganography method.

Nowadays deep learning has achieved impressive results in many fields. Using deep learning in image information hiding is also a hotspot in research. [10] used a deep learning model to recognize the text from images and documents which was further encrypted by an improved encryption method and then was hidden into an image using image steganography. Based on generic encoder-decoder architecture, [11] proposed an automatic steganography method for hiding one image to another. However, it could only embed a unary channel image.

In this paper, we generate the color images with resolution 256×256 which contains secret color images and is similar to the cover images directly, and we can automatically recover the secret color images with resolution 256×256 by the decoder. The framework of blind information hiding based on autoencoder is shown in Figure 1.

Figure 1. The framework of our proposed method to hide the image information based on autoencoder.

Remainder of this paper is organized as follows. Section 2 refers to the application and development course of Autoencoder and deep learning in information hiding. In section 3, we introduce our approach from the formulation and the neural network architecture. Section 4 illustrates the training details and presents the experimental results. In section 5, we further discuss the metrics and defects of our approach.

2. Related works
Autoencoder [12] was a neural network model which used a set of recognition weights to convert an input vector into a code vector, then used a set of generative weights to convert the code vector into an approximate reconstruction of the input vector. Autoencoder was an important deep learning model which could not be ignored in many fields such as image recognition[13], image generation[14], voice recognition[15], image compression[16]. The research of improved method of autoencoder attracted the attention of many scholars. [17] introduced a stochastic variational inference and learning algorithm that scaled to large datasets and, under some mild differentiability conditions, even works in the intractable case. [18] proposed generative adversarial networks (GAN) to perform variational inference by matching the aggregated posterior of the hidden code vector of the autoencoder with an arbitrary prior distribution.
Recently, autoencoder carries out its functions continuously. [19] proposed to utilize deep autoencoder to learn latent representation of high-dimensional mass spectrometry data which was a feasible and powerful instrument for mass spectrometry feature learning and also cancer diagnosis. [20] described a vehicles acoustic event classification algorithm based on sparse autoencoder to analysis the traffic state.

Many methods by Deep learning for image information hiding and steganography had been proposed. Based on Deep Convolutional Generative Adversarial Networks (DCGAN)[21], [22] generated image-like containers and more setanalysis-secure message embedding using Least Significant Bit(LSB)[1] algorithm. [23] used WGAN[24] and GNCNN[25] instead of DCGAN[21] and the discriminator of SGAN[22] to obtain faster convergence speed, better image quality and training stability. [26] proposed an image steganography algorithm and implemented using Multistage Feed-Forward Artificial Neural Network. [27] hided arbitrary binary data in images using generative adversarial networks which optimized the perceptual quality of the images produced by the model.

Recently, [28] proposed three different architectures based on the 3-players game. The first-architecture was proposed as a rigorous alternative to two recent publications. The second took into account stego noise power. Finally, the third architecture enriched the second one with a better interaction between the embedding and extracting networks. Based on image retrieval of DenseNet features and DWT sequence mapping, [29] proposed a novel coverless image steganography which could resist the existing steganalysis tools effectively.

3. Image information hiding based on autoencoder

3.1. formulation
Define x is the secret image, \( x \sim p_{data}(x) \) is the distribution of the secret image \( x \), y is the cover image, encoder G, decoder F. Encoder G generates the encoded image as close as the cover image, decoder has the ability to recover the secret image when input the encoded images. As a result, we propose an encoded loss function as shown in formula (1):

\[
L(G) = E_{x \sim p_{data}(x)} \| G(x) - y \|_2
\]  
(1)

And then we propose a decoded loss function as shown in formula (2):

\[
L(F) = E_{x \sim p_{data}(x)} \| F(G(x)) - x \|_2
\]  
(2)

Finally, we combine the encoded loss function and decoded loss function as our full loss function:

\[
L(G, F) = L(G) + L(F)
\]  
(3)

3.2. Neural network architecture
According to the resolution of secret image and stego image, the structure of encoder and decoder neural network in this paper adopt the generator neural network structure of CycleGAN model[30]. This neural network contains one stride-one convolution and two strides-two convolutions then 9 residual blocks[31] then two strides 1/2 convolutions and one stride-one convolution.

4. Experiments

4.1. training details
In the experiment, Adam solver[32] is selected as the parameter optimizer with a batch size of 1, and the initial learning rate is 0.0002, then decrease to a half after 600 epochs then decrease to a half after 200 epochs then decrease to a half after 100 epochs.
We reshape the image from high resolution DIV2K training set with resolution $256 \times 256$ set as our secret images training set. We resize pictures of Lena with resolution $256 \times 256$ as our cover image.

**Figure 2.** lists two randomly chosen secret images, the cover images, the stego images where the secret image was embedded to the cover image and the reconstructed images which is recovered from the stego images. As can be seen in **Figure 2**, our method can transform a totally different secret image to the picture of Lena, and change it back to the original image automatically.

| Secret image | Cover image | Stego image | Reconstructed image |
|--------------|-------------|-------------|---------------------|
| ![Secret image](image1.png) | ![Cover image](image2.png) | ![Stego image](image3.png) | ![Reconstructed image](image4.png) |

**Figure 2.** Secret images, Cover images, Stego images and Reconstructed images used or obtained by our method.

### 4.2. Hiding and extracting effects analysis

In order to illustrate the hiding effects of our method, we analysis the histogram of cover image and the stego image. As we can see in **Figure 3**, the histograms of the cover image and the stego image in RGB space are almost the same, which proves our method has a good effect in information hiding. In terms of quantitative indicators, we calculate the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) between the cover image and the stego image. As shown in **Table 1**, the indicators further prove the cover image and the stego image is extremely similar, and our method can recover the secret image good enough either.

| | R | G | B |
|---|---|---|---|
| Cover image | ![Histogram](image5.png) | ![Histogram](image6.png) | ![Histogram](image7.png) |
| Stego image | ![Histogram](image8.png) | ![Histogram](image9.png) | ![Histogram](image10.png) |

**Figure 3.** Comparative analysis of the RGB channel decomposition of picture of Lena before and after hiding.
Table 1. The PSNR and SSIM values between the cover image and the stego image, between reconstructed image and secret image.

| Quantized index | between reconstructed image and secret image | between cover image and stego image |
|-----------------|---------------------------------------------|-----------------------------------|
| PSNR            | 41.4026                                     | 39.6868                           |
| SSIM            | 0.9256                                      | 0.9777                            |

4.3. Information hiding capacity analysis

Information hiding capacity refers to the maximum information capacity can be embedded without changing the appearance characteristics of the carriers. The information hiding ability\cite{33} is defined as Equation (4). Table 2 illustrates one pixel of secret image can embed into one pixel of cover image by our method that shows our method has a large hiding capacity.

\[
\text{Relative capacity} = \frac{\text{Absolute capacity}}{\text{The size of the image}} \tag{4}
\]

Table 2. Quantitative analysis of hide capacity of our method

| Absolute capacity (bits/pixel) | Image size | Relative capacity (bits/pixel) |
|--------------------------------|------------|--------------------------------|
| 256×256                        | 256×256    | 1                              |

5. Conclusion

In order to hide image information automatically and improve the steganography capacity, in this paper, we introduce a method of blind information hiding method based on the autoencoder. The encoder embeds the secret image into the cover image and the decoder recover the secret image from stego image. Experimental results show our method has a good ability in image information hiding and steganography capacity. In the future, the resistance ability to multi-attacks of this method should be further considered.

Acknowledgments

This work was financially supported in part by the National Key Research and Development Project of China under Grant 2017YFB1402100, in part by the Natural Science Foundation of Xizang Autonomous Region of China under Grant XZ2018ZR G-64.

References

[1] Mielikainen, Jarno. Lsb matching revisited. IEEE signal processing letters, 13(5):285–287, 2006.
[2] Zhou, Jia Min, et al. “DCT-Based Digital Image Steganography.” Applied Mechanics and Materials, vol. 496–500, Trans Tech Publications, Ltd., Jan. 2014, pp. 1986–1990. Crossref, doi:10.4028/www.scientific.net/amm.496-500.1986.
[3] Sheikhan, Mansour , K. Asadollahi , and R. Shahnazi . "Improvement of Embedding Capacity and Quality of DWT-Based Audio Steganography Systems. "World Applied Sciences Journal"13.3(2013).
[4] Luo, Weiqi, F. Huang, and J. Huang. "Edge Adaptive Image Steganography Based on LSB Matching Revisited." IEEE Transactions on Information Forensics & Security5.2(2010):201-214.
[5] Boukis, Andreas C., et al. "Multicomponent reactions provide key molecules for secret communication." Nature Communications 9.1(2018).
[6] Zhang, Chenggong , et al. "Compressive optical steganography via single-pixel imaging." Optics Express27.9(2019):13469.
[7] Akgul, Akif , I. M. Moroz , and A. Durdu . "A novel data hiding method by using a chaotic
system without equilibrium points." Modern Physics Letters B 33.1(2019).

[8] Nandhini, et al. "High capacity multi-bit data hiding based on modified histogram shifting technique." Etri Journal (2018).

[9] Karakus, Songul, and E. Avci. "A New Image Steganography Method with Optimum Pixel Similarity for Data Hiding in Medical Images." Medical Hypotheses 139(2020):109691.

[10] Sathiaraj, S. J. Fiona G., et al. "Secure Transfer of Image-Acquired Text Using a Combination of Cryptography and Steganography." 2019 1st International Conference on Advances in Information Technology (ICAIT) IEEE, 2019.

[11] Rehman, Atique & Rahim, Rafia & Nadeem, Muhammad & Hussain, Sibt. (2017). End-to-end Trained CNN Encode-Decoder Networks for Image Steganography.

[12] Hinton, G. E. . "Autoencoders, minimum description length and Helmholtz free energy." Advances in Neural Information Processing Systems San Mateo 6(1994).

[13] Shi, Lukui, C. Gao, and J. Zhang. . "Pavement Distress Image Recognition Based on Multilayer Autoencoders." International Conference on Artificial Intelligence & Computational Intelligence Springer Berlin Heidelberg, 2012.

[14] Khan, Salman & Hayat, Munawar & Barnes, Nick. (2018). Adversarial Training of Variational Auto-Encoders for High Fidelity Image Generation. 1312-1320. 10.1109/WACV.2018.00148.

[15] Mimura, Masato, S. Sakai, and T. Kawahara. . "Exploring deep neural networks and deep autoencoders in reverberant speech recognition." Hands-free Speech Communication & Microphone Arrays IEEE, 2014.

[16] Zebang, Song & Sei-ichiro, Kamata. (2019). Densely connected AutoEncoders for image compression. ICIIGP '19: Proceedings of the 2nd International Conference on Image and Graphics Processing. 78-83. 10.1145/3313950.3313965.

[17] Kingma, Diederik P, and M. Welling. "Auto-Encoding Variational Bayes." (2013).

[18] Makhzani, Alireza, et al. "Adversarial Autoencoders." Computer Science (2015).

[19] Zhou, Qingguo, et al. "Deep Autoencoder for Mass Spectrometry Feature Learning and Cancer Detection." IEEE Access 8.99(2020):45156-45166.

[20] Zhang, Xiaodan, Y. Chen, and G. Tang. . "Research on Traffic Acoustic Event Detection Algorithm Based on Sparse Autoencoder." MATEC Web of Conferences 308.7(2020):05002.

[21] Radford, Alec, L. Metz, and S. Chintala. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks." Computer Science (2015).

[22] Denis Volkonskiy, Boris Borisenko and Burnaev, Evgeny. "Generative adversarial networks for image steganography." In ICLR 2016 Open Review.

[23] Shih H., Dong J., Wang W., Qian Y., Zhang X. "SSGAN: Secure Steganography Based on Generative Adversarial Networks." (2017).

[24] Arjovsky, M., Chintala, S., and Bottou, L. Wasserstein GAN. ArXiv e-prints, January 2017.

[25] Zhu, Jun Yan, et al. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks." (2017).
[31] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016.
[32] D. Kingma and J. Ba. Adam: A method for stochastic optimization. In ICLR, 2015.
[33] Duan, Xintao , et al. "High-Capacity Information Hiding Based on Residual Network." IETE Technical Review 4(2020):1-12.