 JIHPP, vol.1, no.2, pp.87-93, 2019

A Survey on Digital Image Steganography

Jiaxin Wang[1, *], Mengxin Cheng[1], Peng Wu[1] and Beijing Chen[1, 2]

Abstract: Internet brings us not only the convenience of communication but also some security risks, such as intercepting information and stealing information. Therefore, some important information needs to be hidden during communication. Steganography is the most common information hiding technology. This paper provides a literature review on digital image steganography. The existing steganography algorithms are classified into traditional algorithms and deep learning-based algorithms. Moreover, their advantages and weaknesses are pointed out. Finally, further research directions are discussed.

Keywords: Information hiding, image steganography, information security, deep learning.

1 Introduction

With the rapid development of information processing technology, computer technology and network communication technology, human has entered the digital information age, and the Internet has become a new channel for transmitting information. However, Some security risks have also appeared, such as intercepting information and stealing information. Therefore, ensuring the security of information on the Internet has become one of the problems to be solved urgently [Gupta, Gupta and Singhal (2014)].

In order to ensure the security of information transmitted on the Internet and prevent information leakage, information hiding technology has emerged. Information hiding technology can be divided into two important branches: watermarking and steganography. Digital watermarking hides secret information in digital media (such as pdf documents, word documents, videos, etc.) to prevent piracy and protect copyright of the original. Steganography is mainly used for covert communication, so that the secret information hidden in the carrier data does not attract the attention of external observers. Generally speaking, steganography implements a “peer-to-peer” covert communication. Only the communicating parties know the existence of the secret message, and the receiver can accurately extract the secret information. Steganography is a very important tool in the protection of national information security, especially in the military. Therefore, steganography is a very important research content in the field of information security.

Most of image steganography algorithms can be divided into two stages (As shown in Fig.

[1] School of Computer & Software, Nanjing University of Information Science & Technology, Nanjing, 210044, China.

[2] Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Nanjing University of Information Science & Technology, Nanjing, 210044, China.

[*] Corresponding Author: Jiaxin Wang. Email: jhx1712184106@163.com.

1): 1) Embedding. In this stage, secret message is embedded into carrier image; 2) Extracting. It realizes the function of recovering secret message from steganographic images.

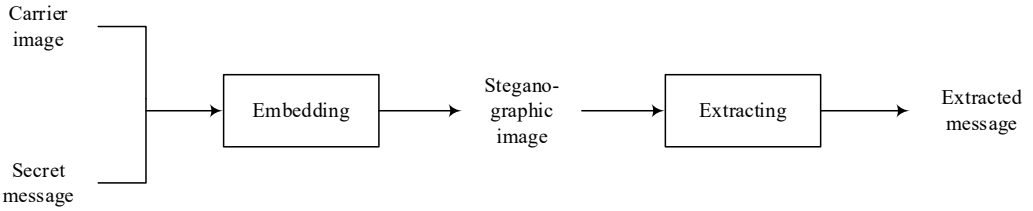

Figure 1: Steps in steganography

In this paper, some traditional image steganography algorithms and recent algorithms based on deep learning are surveyed. The rest of the paper is organized as follows: Section 2 reviews some traditional algorithms from spatial-domain and transform-domain. Section 3 describes the deep learning-based algorithms. Future research directions are discussed in Section 4. Section 5 summarizes the paper.

2 Traditional algorithms

In recent years, many researchers have studied and explored steganography techniques and provided many excellent steganography algorithms. From the perspective of embedding domain, traditional steganography algorithms can be divided into spatial-domain algorithms and transform-domain algorithms [Carvajal-Gamez, Gallegos-Funes and Rosales-Silva (2013)].

2.1 Spatial-domain algorithms

The spatial-domain algorithms mainly hide secret information by changing the brightness value or chrominance value of carrier image. Least significant bit (LSB) algorithm is a simple and representative spatial-domain algorithm. It first converts secret information into binary data, then replaces the LSB of some pixels of the carrier image [Bender, Gruhl, Morimoto et al. (1996); Lie and Chang (1999)]. The LSB algorithm has low complexity and high embedding capacity, and is easy to implement. However, its robustness against steganalysis is not very strong. Therefore, many improved LSB algorithms have been proposed to solve this problem. Chan et al. [Chan and Cheng (2004)] proposed a simple and effective algorithm called Optimal Pixel Adjustment Process (OPAP) to reduce the distortion caused by LSB. Later, Kekre et al. [Kekre, Athawale and Athawale (2011)] proposed an improved algorithm called Multiple Least Significant Bit (MLSB) to replace the lowest 3 significant bits of the carrier image with secret information. Compared with the original LSB algorithm, the MLSB algorithm has a higher payload. Recently, Sarreshtedari et al. [Sarreshtedari and Akhaee (2013)] improved the invisibility of steganographic image by reducing the probability of modifying pixels in carrier image, moreover their proposed algorithm can resist the HCF-COM Steganalysis. Tavares et al. [Tavares and Junior (2016)] reduced the modification of pixel value when hiding information, thus indirectly increased visual quality of the steganographic image. Moreover, some works contribute to improving the security of LSB algorithm by combining some different encryption algorithms [Amirtharajan and

Rayappan (2012); Muhammad, Ahmad, Rehman et al. (2017)].

2.2 Transform-domain algorithms

Different from the spatial-domain steganography algorithms, the transform-domain algorithms embed secret information in the transform domain of the carrier image under different kinds of transforms. The transforms include Discrete Wavelet Transform (DWT) [Chen (2007); Abdelwahab and Hassaan (2008); Su and Kuo (2003); Zhang, Wang and Wu (2009); Chen and Lin (2006)], Discrete Fourier Transform (DFT) [Ramkumar, Akansu and Alatan (1999); Rivas (2007); Khashandarag and Ebrahimian (2009)], Discrete Cosine Transform (DCT) [Chang, Chen and Chung (2002); Liu, Dai, Sun et al. (2007); Quan, Hua and Zu (2009); Almohammad, Hierons and Ghinea (2008); Vongurai and Phimoltares (2012); Banu and Shajeesh (2017)] and so on. Among them, the DCT is the most widely used one because it is compatible with JPEG compression standards.

The DCT-based algorithms mainly modify the quantized DCT coefficients to embed secret information. They are based on a custom quantization table. Chang et al. [Chang, Chen and Chung (2002)] proposed a JPEG steganography algorithm based on 8×8 quantization table, which improves the embedding capacity of secret information while ensuring image visual quality. Liu et al. [Liu, Dai, Sun et al. (2007)] proposed an improved steganography algorithm based on quantization tables and the steepest descent method, which guarantees the steganographic image with a low distortion. Quan et al. [Quan, Hua and Zu (2009)] presented a large-capacity steganography algorithm based on DCT but greatly increasing the complexity. In addition, some researchers increased the size of custom quantization table to increase the capacity. For example, Almohammad et al. [Almohammad, Hierons and Ghinea (2008)] used 16×16 quantization table, while Vongurai et al. [Vongurai and Phimoltares (2012)] used an extended 32×32 quantization table. Banu et al. [Banu and Shajeesh (2017)] made use of an interpolation quantization table.

In summary, the spatial-domain steganography algorithm has the advantages of large capacity, good visual quality and easy implementation. Due to these advantages of the spatial-domain steganography algorithm, many steganography algorithms borrow or use spatial-domain steganography or its derived algorithms. Therefore, this kind of algorithm has become the most representative method in digital image steganography. However, spatial-domain algorithms have many advantages, which also have some drawbacks. For example, spatial-domain algorithms are easily detected by corresponding methods. Transform-domain algorithms tend to have better anti-attack capabilities than spatial-domain algorithms, but in general, they have less payload.

3 Deep learning-based algorithms

In recent years, the development of deep learning has injected new vitality into the field of steganography. Some researchers have been working on improving deep learning network structure with the consideration of the characteristics of steganography.

The steganography algorithms based on deep learning are basically realized by adversarial network. Volkhonskiy et al. [Volkhonskiy, Nazarov, Borisenko et al. (2017)] first proposed the steganography model called Steganographic Generative Adversarial Network (SGAN) based on GAN. This model resisted steganalysis and made hidden

information secure. On the basis of SGAN, Shi et al. [Shi, Dong, Wang et al. (2017)] proposed SSGAN (Secure Steganography Based on GAN) to enhance the security against steganalysis. The HayesGAN model proposed by Hayes et al. [Hayes and Danezis (2017)] used adversarial learning to generate steganographic images directly. Then, Hu et al. [Hu, Wang, Jiang et al. (2018)] introduced a steganalysis network into HayesGAN to improve the quality and safety of the generated steganographic images. However, these two works are not guaranteed to completely extract the embedded secret information. Zhu et al. [Zhu, Kaplan, Johnson et al. (2018)] proposed another model called Hiding Data with Deep Network (HiDDeN) based on HayesGAN. It can extract embedded information with high accuracy though the existence of various attacks, such as Gaussian blur, missing pixels, cropping, and JPEG compression, etc. Tang et al. [Tang, Tan, Li et al. (2017)] combined the GAN with adaptive steganography algorithm to find suitable steganographic positions for steganography and proposed Automatic Steganographic Distortion Learning Framework with GAN (ASDL-GAN). Yang et al. [Yang, Liu, Kang et al. (2018)] modified the ASDL-GAN model through replacing the activation function Ternary Embedding Simulator (TES) by Tanh to improve security.

Although the above-mentioned adversarial network-based steganography algorithms achieve fine performance in steganography and also in resisting steganalysis, their hiding capacities are very limited. So, some researchers worked on large-capacity steganography algorithms to embed secret images into carrier images with the same size. Rahim et al. [Rahim and Nadeem (2017)] proposed an end-to-end framework to embed a secret gray image into a color carrier image with the same size. Their work realizes high-capacity embedding but distorts the steganographic image in color. Then, in the loss function Baluja et al. [Baluja (2017)] considered the correlation between secret image and error image obtaining from steganographic image and carrier image to improve the invisibility of the steganographic image. Zhang et al. [Zhang, Dong and Liu (2018)] proposed ISGAN (Invisible Steganography via Generative Adversarial Networks) by introducing the steganalysis network proposed by Xu et al. [Xu, Wu and Shi (2016)] into their basic model to improve its ability to resist steganalysis. In addition, the SteganoGAN model proposed by Zhang et al. [Zhang, Cuesta-Infante, Xu et al. (2019)] used residual structure to further improve the quality of steganographic image.

4 Future research directions

The existing algorithms presented in Section 3 have achieved fine performance in steganography, but there still remain some unresolved issues.

4.1 How to achieve a balance between security and capacity

For both of traditional algorithms and deep learning-based algorithms, there is such a problem: with the increase of steganography capacity, steganography algorithm is more easily detected by steganalysis. Then, some algorithms reduce the capacity to resist the steganalysis. So, how to achieve a balance between capacity and security? It needs to be explored in future work.

4.2 How to improve the quality of steganographic image from the deep learning-based large-capacity steganography algorithm

For the deep learning-based large-capacity steganography algorithm, the quality of both of steganographic images and extracted images need to be improved. For example, in the work of Rahim et al. [Rahim and Nadeem (2017)], there is a color distortion of the steganographic image. Such color distortion is not feasible in the steganography algorithm. Therefore, the quality of steganographic image should be paid more attention.

4.3 How to consider robustness

Most of steganography algorithms mainly consider the ability of hiding information and also resisting steganalysis, while they ignore the consideration of robustness. So, the algorithms are not robust. In future work, for the machine learning-based algorithm, the robustness should be considered in both of training phase and testing phase.

5 Conclusion

This paper starts with introducing the background of image steganography. Then, the state-of-the-art algorithms are summarized from spatial-domain algorithms, transform-domain algorithms to deep learning-based algorithms. What's more, the current existing issues and future research directions are also illustrated in this paper.

Acknowledgement: This work was supported by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant 17KJB520021, the PAPD fund, sponsored by Qing Lan Project.

References

Abdelwahab, A. A.; Hassaan, L. A. (2008): A discrete wavelet transform based technique for image data hiding. *Proceedings of 2008 National Radio Science Conference*, pp. 1-9.

Almohammad, A.; Hierons, R. M.; Ghinea, G. (2008): High capacity steganographic method based upon JPEG. *Proceedings of 2008 Third International Conference on Availability, Reliability and Security*, pp. 544-549.

Amirtharajan, R.; Rayappan, J. B. B. (2012): An intelligent chaotic embedding approach to enhance stego-image quality. *Information Sciences*, vol. 193, pp. 115-124.

Baluja, S. (2017): Hiding images in plain sight: deep steganography. *Advances in Neural Information Processing Systems*, pp. 2069-2079.

Banu, T. S.; Shajeesh, K. U. (2017): Secure reversible data hiding technique on textures using double encryption. *Proceedings of 2017 International Conference on Innovations in Information, Embedded and Communication Systems*, pp. 1-5.

Bender, W.; Gruhl, D., Morimoto, N.; Lu, A. (1996): Techniques for data hiding. *IBM Systems Journal*, vol. 35, no. 3, pp. 313-336.

Carvajal-Gamez, B. E.; Gallegos-Funes, F. J.; Rosales-Silva, A. J. (2013): Color local complexity estimation based steganographic (CLCES) method. *Expert Systems with Applications*, vol. 40, no. 4, pp. 1132-1142.

Chan, C. K.; Cheng, L. M. (2004): Hiding data in images by simple LSB substitution. *Pattern Recognition*, vol. 37, no. 3, pp. 469-474.

Chang, C. C.; Chen, T. S.; Chung, L. Z. (2002): A steganographic method based upon JPEG and quantization table modification. *Information Sciences*, vol. 141, no. 1, pp. 123-138.

Chen, P. Y.; Lin, H. J. (2006): A DWT based approach for image steganography. *International Journal of Applied Science and Engineering*, vol. 4, no. 3, pp. 275-290.

Chen, W. Y. (2007): Color image steganography scheme using set partitioning in hierarchical trees coding, digital Fourier transform and adaptive phase modulation. *Applied Mathematics and Computation*, vol. 185, no. 1, pp. 432-448.

Gupta, R.; Gupta, S.; Singhal, A. (2014): Importance and techniques of information hiding: a review. *International Journal of Computer Trends & Technology*, vol. 9, no. 5.

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

Hu, D.; Wang, L.; Jiang, W.; Zheng, S.; Li, B. (2018): A novel image steganography method via deep convolutional generative adversarial networks. *IEEE Access*, vol. 6, pp. 38303-38314.

Kekre, H. B.; Athawale, A. A.; Athawale, U. A. (2011): Increased cover capacity using advanced multiple LSB algorithms. *International Conference & Workshop on Emerging Trends in Technology*, pp. 25-31.

Khashandarag, A. S.; Ebrahimian, N. (2009): A new method for color image steganography using SPIHT and DFT, sending with JPEG format. *Proceedings of 2009 International Conference on Computer Technology and Development*, vol. 1, pp. 581-586.

Lie, W. N.; Chang, L. C. (1999): Data hiding in images with adaptive numbers of least significant bits based on the human visual system. *Proceedings of 1999 International Conference on Image Processing*, vol. 1, pp. 286-290.

Liu, G. J.; Dai, Y. W.; Sun, J. S.; Wang, Z. Q. (2007): High capacity information hiding scheme for JPEG images. *Information and Control*, vol. 36, no. 1, pp. 102-107.

Muhammad, K.; Ahmad, J.; Rehman, N. U.; Jan, Z.; Sajjad, M. (2017): CISSKA-LSB: color image steganography using stego key-directed adaptive LSB substitution method. *Multimedia Tools and Applications*, vol. 76, no. 6, pp. 8597-8626.

Quan, X. J.; Hua, Y. C.; Zu, H. D. (2009): A high capacity information hiding algorithm in DCT domain. *Journal of Image and Graphics*, vol. 14, no. 8, pp. 1542-1546.

Rahim, R.; Nadeem, M. S. (2017): End-to-end trained CNN encode-decoder networks for image steganography. *European Conference on Computer Vision*, pp. 723-729.

Ramkumar, M.; Akansu, A. N.; Alatan, A. A. (1999): A robust data hiding scheme for images using DFT. *Proceedings of 1999 International Conference on Image Processing*, vol. 2, pp. 211-215.

Rivas, E. (2007): Fourier phase domain steganography: phase bin encoding via interpolation. *Proceedings of SPIE-the International Society for Optical Engineering*, vol. 6579.

Sarreshtedari, S.; Akhaee, M. A. (2013): One-third probability embedding: a new±1 histogram compensating image least significant bit steganography scheme. *IET Image Processing*, vol. 8, no. 2, pp. 78-89.

Shi, H.; Dong, J.; Wang, W.; Qian, Y.; Zhang, X. (2017): SSGAN: secure steganography based on generative adversarial networks. *Pacific Rim Conference on Multimedia*, pp. 534-544.

Su, P. C.; Kuo, C. C. (2003): Steganography in JPEG2000 compressed images. *IEEE Transactions on Consumer Electronics*, vol. 49, no. 4, pp. 824-832.

Tang, W.; Tan, S.; Li, B.; Huang, J. (2017): Automatic steganographic distortion learning using a generative adversarial network. *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1547-1551.

Tavares, J. R. C.; Junior, F. M. B. (2016): Word-Hunt: a LSB steganography method with low expected number of modifications per pixel. *IEEE Latin America Transactions*, vol. 14, no. 2, pp. 1058-1064.

Volkhonskiy, D.; Nazarov, I.; Borisenko, B.; Burnaev, E. (2017): Steganographic generative adversarial networks.

Vongurai, N.; Phimoltares, S. (2012): Frequency-based steganography using 32×32 interpolated quantization table and discrete cosine transform. *Fourth International Conference on Computational Intelligence, Modelling and Simulation*, pp. 249-253.

Xu, G.; Wu, H. Z.; Shi, Y. Q. (2016): Structural design of convolutional neural networks for steganalysis. *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 708-712.

Yang, J.; Liu, K.; Kang, X.; Wong, E. K.; Shi, Y. Q. (2018): Spatial image steganography based on generative adversarial network. arXiv:1804.07939.

Zhang, K. A.; Cuesta-Infante, A.; Xu, L.; Veeramachaneni, K. (2019): SteganoGAN: high capacity image steganography with gans. arXiv:1901.03892.

Zhang, L.; Wang, H.; Wu, R. (2009): A high-capacity steganography scheme for JPEG2000 baseline system. *IEEE transactions on Image Processing*, vol. 18, no. 8, pp. 1797-1803.

Zhang, R.; Dong, S.; Liu, J. (2018): Invisible steganography via generative adversarial networks. *Multimedia Tools and Applications*, pp. 1-17.

Zhu, J.; Kaplan, R.; Johnson, J.; Li, F. F. (2018): Hidden: hiding data with deep networks. *Proceedings of the European Conference on Computer Vision*, pp. 657-672.