Detection of Image Steganography

Subjects: Computer Science, Artificial Intelligence

Contributor: Dina Yousif Mikhail , Roojwan Sc Hawezi , Shahab Wahhab Kareem

As internet traffic grows daily, so does the need to protect it. Network security protects data from unauthorized access and ensures their confidentiality and integrity. Steganography is the practice and study of concealing communications by inserting them into seemingly unrelated data streams (cover media). Investigating and adapting machine learning models in digital image steganalysis is becoming more popular.

deep learning transfer learning steganography feature extraction

1. Introduction

As the quantity of traffic being moved and communicated over the internet in various formats, such as movies and photographs, increases daily, there is a rising concern about the security of the massive amount of data transferred over the internet, including passwords and personal, professional, and financial information, as well as social security numbers and other sensitive data. The need to keep this information safe is rising as well. Network security has become an integral part of today's sophisticated communication infrastructure to keep information private and prevent tampering [1][2].

Using machine learning to analyze and improve digital images, the practice of steganalysis is gaining in popularity. Steganography methods implemented inside such a framework are more secure than those employing manually constructed pricing. Steganalysis refers to the process of identifying hidden messages using steganography. In the world of cryptography, this is analogous to the practice of cryptanalysis. As a result, learning stenography is essential. A communication that has been intercepted and decrypted is present whereas cryptanalysis begins with a collection of suspicious data files without knowing which files, if any, hold a payload. The first step in steganalysis, which effectively requires a forensic statistician, is to identify the data files that are most likely to have been tampered with in a massive, often exhaustive, collection [3].

There is a constant, growing requirement to safeguard the escalating volume of sensitive information exchanged and communicated online daily in various formats, including movies and photographs. The need for secure networks to protect sensitive data from prying eyes and malicious actors has made such security an integral aspect of today's communication infrastructure [4][5][6][7][8].

Cryptography, which is the practice of transforming plain text into encrypted text via an algorithm, has been widely used for decades to safeguard sensitive information. To read an original message, a recipient must first convert encrypted text into plain text (Smid and Branstad, 1988) [4]. To keep sensitive data safe, encryption techniques

such as the advanced encryption standard (AES) (NIST-FIPS, 2001) and the data encryption standard (DES) (NIST-FIPS, 1977) are commonly utilized (Yegireddi and Kumar, 2016) [5]. The fact that encrypted communication can be read by anyone is seen as a weakness in cryptography, which translates secret messages into human-readable forms. Therefore, hackers on the internet may use the heat, tries, and attempts strategy to decipher code. Because of this shortcoming in cryptography, steganography was brought into the field of data protection to circumvent this problem by disguising the fact that a communication was taking place.

2. Research on Steganography Detection

Article [3] discussed how to combine trained CDNs in a multimodal framework, and it examined their detection accuracy. The framework detects each classification modality independently and combines their estimations to create a universal image steganography detector. Six of the latest CDN-based image steganography detection techniques—GNCNN, IGNCNN, XuNet, YeNet, YedroudjNet, and the improved IGNCNN—were trained on stego images generated using WOW, SUNIWARED, and HILL steganography algorithms with payloads of 0.2, 0.3, and 0.4 bits per pixel. Due to the projected similarities between the image steganography systems, the detection accuracy decreased slightly. However, the multimodal image steganography detection based on the improved IGNCNN universal image steganography detection performed best compared to the other five examined detectors [3].

Article [9] discussed detecting steganography-modified JPEG images, and it analyzed image steganography detection using shallow and deep learning methods. Three common stenographic algorithms—JPEG universal wavelet relative distortion (J-Uniward), nsF5, and uniform embedding revisited distortion (UERD) at two density levels—hid information in BOSS database photos. DCTR and GFR were the best feature spaces validated. At 0.4, the nsF5 algorithm detected bpnzac density with 99.9% accuracy, but the J-Uniward algorithm was barely detectable at 0.1 (a maximum of 56.3%). The study concluded that ensemble classifiers were a promising alternative to deep-learning-based detection [9].

In [10], the authors presented a deep-learning-based approach for steganography detection in digital images. The authors began by describing the importance of steganography detection in the field of digital forensics and highlighted the challenges associated with it. They then proposed a CNN-based model for detecting the presence of hidden data in digital images. The proposed model took as input the pixel values of an image and learned to identify the presence of steganography through a series of convolutional, pooling, and fully connected layers. The authors evaluated the performance of the proposed model on a dataset of stego images and showed that it outperformed existing steganography detection techniques in terms of accuracy, precision, and recall. The results of the study suggested that deep-learning-based approaches can be effective for steganography detection in digital images and can help improve the accuracy and reliability of forensic investigations.

In [11], the authors began by describing the importance of steganalysis in digital forensics and highlighted the limitations of traditional steganalysis techniques. They then proposed a deep-learning-based model for steganalysis that used a combination of convolutional and fully connected neural networks. The proposed model

was trained and evaluated on a dataset of stego images that contained spatially embedded hidden information. The authors showed that the proposed model outperformed existing steganalysis techniques in terms of accuracy and sensitivity to different types of spatial image steganography. They also performed a sensitivity analysis of the proposed model to evaluate the impacts of different hyper parameters and architecture choices on the model's performance. The results of the study suggested that deep-learning-based approaches can be highly effective for the steganalysis of spatially embedded hidden information in digital images, and that careful selection of hyper parameters and architecture choices can further improve the performance of a model [11].

The authors of [12] began by describing the importance of steganalysis in digital forensics and highlighted the limitations of traditional steganalysis techniques. They then proposed a deep-learning-based model for steganalysis that used a combination of non-local blocks and multi-channel convolutional networks to identify the presence of hidden information in an image. The proposed model was trained and evaluated on a dataset of stego images that contained spatially embedded hidden information. The authors showed that the proposed model outperformed existing steganalysis techniques in terms of accuracy, precision, and recall. They also showed that the proposed model could be used to localize the regions of an image that contained hidden information. The results of the study suggested that deep-learning-based approaches can be highly effective for the steganalysis of spatially embedded hidden information in digital images and that the proposed model can help improve the accuracy and efficiency of forensic investigations [12].

Article [13] presented a deep-learning-based approach for detecting steganography in color images. The authors began by describing the importance of steganalysis in digital forensics and highlighted the challenges associated with detecting hidden information in color images. They then proposed a multi-frequency residual convolutional neural network (MRF-CNN) for steganalysis that extracted features from different frequency components of an image and learned to identify the presence of hidden information. The proposed model was trained and evaluated on a dataset of stego color images and compared with existing steganalysis techniques. The authors showed that the proposed MRF-CNN model outperformed existing steganalysis techniques in terms of accuracy, precision, and recall. They also showed that the proposed model could be used to localize the regions of an image that contained hidden information. The results of the study suggested that deep-learning-based approaches, such as the proposed MRF-CNN model, can be highly effective for the steg analysis of images and can help improve the accuracy and efficiency of forensic investigations.

The authors of [14] began by describing the importance of hand movement identification in various applications, such as prosthetics and rehabilitation. They then proposed a machine-learning-based approach for identifying hand movements that involved the use of electromyography (EMG) signals recorded from muscles in the arm. The proposed approach used a combination of feature extraction techniques and classification algorithms to identify hand movements. The authors evaluated the proposed approach on a dataset of EMG signals recorded from multiple subjects and showed that the proposed approach achieved high accuracy in identifying hand movements. They also compared the performance of the proposed approach with that of existing approaches and showed that the proposed approach outperformed them. The results of the study suggested that machine-learning-based

approaches can be highly effective for identifying hand movements and can have important applications in prosthetics and rehabilitation [14].

In [15], the authors presented a novel deep neural network for identifying contextual steganography methods. The suggested method employed a high-boost filter to reduce high-frequency noise while keeping low-frequency information intact. Thirty high-pass SRM filters were applied to the high-boost image, resulting in thirty high-boost SRM-filtered photos. The suggested CNN used two skip connections to simultaneously gather data from a large number of connections. Despite the standard ReLU layer, a cropped version was investigated. The convolutional neural network (CNN) was built using a bottleneck strategy for maximum efficiency. For comprehensive data persistence, only one layer of global average pooling was used. To further enhance the detection accuracy, SVM was used in place of the softmax classifier. Compared to state-of-the-art methods, the proposed method performed better in terms of detection accuracy and computational cost in the experiments. The HILL, S-UNIWARD, and WOW context-aware steganography algorithms were tested on the BOWS2 and BOSS base datasets, validating the suggested scheme [15].

References

1. Butora, J.; Yousfi, Y.; Fridrich, J. How to pretrain for steganalysis. In Proceedings of the 2021 ACM Workshop on Information Hiding and Multimedia Security, Virtual, 21–25 June 2021; pp. 143–148.
2. Chaumont, M. Deep learning in steganography and steganalysis. In Digital Media Steganography; Academic Press: Cambridge, MA, USA, 2020; pp. 321–349.
3. Elshafey, M.A.; Amein, A.S.; Badran, K.S. Universal Image Steganography Detection using Multimodal Deep Learning Framework. J. Inf. Hiding Multim. Signal Process. 2021, 12, 152–161.
4. Smid, M.E.; Branstad, D.K. Data Encryption Standard: Past and future. Proc. IEEE 1988, 76, 550–559.
5. Yegireddi, R.; Kumar, R.K. A survey on conventional encryption algorithms of Cryptography. In Proceedings of the 2016 International Conference on ICT in Business Industry & Government (ICTBIG), Indore, India, 18–19 November 2016; pp. 1–4.
6. Ozcan, S.; Mustacoglu, A.F. Transfer learning effects on image steganalysis with pre-trained deep residual neural network model. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 2280–2287.
7. Reinel, T.-S.; Brayan, A.-A.H.; Alejandro, B.-O.M.; Alejandro, M.-R.; Daniel, A.-G.; Alejandro, A.-G.J.; Buenaventura, B.-J.A.; Simon, O.-A.; Gustavo, I.; Raul, R.-P. GBRAS-Net: A convolutional neural network architecture for spatial image steganalysis. IEEE Access 2021, 9, 14340–14350.

8. Selvaraj, A.; Ezhilarasan, A.; Wellington, S.L.J.; Sam, A.R. Digital image steganalysis: A survey on paradigm shift from machine learning to deep learning based techniques. IET Image Process. 2021, 15, 504–522.
9. Płachta, M.; Krzemień, M.; Szczypiorski, K.; Janicki, A. Detection of Image Steganography Using Deep Learning and Ensemble Classifiers. Electronics 2022, 11, 1565.
10. Reinel, T.S.; Raul, R.P.; Gustavo, I. Deep Learning Applied to Steganalysis of Digital Images: A Systematic Review. IEEE Access 2019, 7, 68970–68990.
11. Tabares-Soto, R.; Arteaga-Arteaga, H.B.; Mora-Rubio, A.; Bravo-Ortíz, M.A.; Arias-Garzón, D.; Alzate-Grisales, J.A.; Orozco-Arias, S.; Isaza, G.; Ramos-Pollán, R. Sensitivity of deep learning applied to spatial image steganalysis. PeerJ Comput. Sci. 2021, 7, e616.
12. Han, X.; Zhang, T. Spatial Steganalysis Based on Non-Local Block and Multi-Channel Convolutional Networks. IEEE Access 2022, 10, 87241–87253.
13. Lin, J.; Yang, Y. Multi-Frequency Residual Convolutional Neural Network for Steganalysis of Color Images. IEEE Access 2021, 9, 141938–141950.
14. Mora-Rubio, A.; Alzate-Grisales, J.A.; Arias-Garzón, D.; Buriticá, J.I.P.; Varón, C.F.J.; Bravo-Ortiz, M.A.; Arteaga-Arteaga, H.B.; Hassaballah, M.; Orozco-Arias, S.; Isaza, G.; et al. Multi-subject identification of hand movements using machine learning. In Sustainable Smart Cities and Territories; Corchado, J.M., Trabelsi, S., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 117–128.
15. Agarwal, S.; Kim, C.; Jung, K.-H. Steganalysis of Context-Aware Image Steganography Techniques Using Convolutional Neural Network. Appl. Sci. 2022, 12, 10793.