*International Journal for Modern Trends in Science and Technology*, 9(02): 151-154, 2023

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0902027
Available online at: http://www.ijmtst.com/vol9issue02.html

# Convolutional Neural Network Based Text Steganalysis

**Ch.Santha Kumari [1] | S. Naga Niharika [2] | T. Sethu Vinay [2] | Sk. Reshma [2]| P. Likitha [2]**

[1]Assistant Professor, Department of CSE, NRI Institute of Technology, India
[2]B.Tech Student, Department of CSE, NRI Institute of Technology, India

**To Cite this Article**
Ch.Santha Kumari, S. Naga Niharika, T. Sethu Vinay, Sk. Reshma, P. Likitha. Research on Convolutional Neural Network Based Text Steganalysis. International Journal for Modern Trends in Science and Technology 2023, 9(02), pp. 151-154. https://doi.org/10.46501/IJMTST0902027

**Article Info**
Received: 30 December 2022; Accepted: 01 February 2023; Published: 04 February 2023

## ABSTRACT

***The current state of the art in text steganalysis involves extracting human-crafted characteristics and categorizing them using support vector machines to identify steganographic communication. Due to the fact that they are developed on the basis of statistical modifications brought about by steganography, these characteristics are not easily transferable to other embedding techniques, and their detection efficiency is very size-dependent. In this letter, we offer a new convolutional neural network-based model for text steganalysis that can automatically learn feature representations from texts and capture complicated relationships. To begin, we use a word embedding layer to glean linguistic information. As a second step, we utilize varying-sized rectangular convolution kernels to learn grammatical characteristics of sentences. We also provide a judgment approach for identifying these lengthy texts in an effort to boost performance even more. The experimental findings demonstrate the effectiveness of the proposed technique in detecting various text steganographic algorithms, with results that are either on par with or above those of the state-of-the-art methods over a broad range of text sizes.***

***KEY WORDS: Text steganalysis, convolutional neural network, deep learning, decision strategy***

## 1. INTRODUCTION

Text steganography is a form of steganography that makes use of natural language processing in order to embed secret data into texts. This form of steganography has garnered a lot of attention in recent years due to the prevalence of texts (books, emails, blogs, etc.) that can serve as rich carriers for steganography.

Text steganalysis is the art and science of preventing the covert communication between criminal offenders. It may be thought of as the antidote to the practice of steganography, which hides communications with the intention of revealing their presence.

The majority of the text steganalysis algorithms [1]–[6] are built on a generic machine learning framework. This framework involves the extraction of a collection of handmade features to capture the artifacts of embedding procedures, which are then given to a classifier such as an SVM. However, due to the fact that diverse steganography methods need uniquely built features, these characteristics do not have general applicability. Additionally, in order to acquire characteristics that are more useful, it is necessary to make use of the information that is specific to the domain in question. Furthermore, the features are dependent on the statistical characteristics that must be well manifested in long texts. Because of this, the

performance of the existing text steganalysis schemes is better for long texts, but it is worse for short texts (within 200 words). This is despite the fact that many texts are quite short, such as tweets, e-mails, letters, and movie reviews.

The accomplishments of CNN-based picture steganography [7–9] have motivated us to develop a text steganography model that is built on a deep learning architecture. This model will use short sentences as its input and will differentiate between stego texts and cover texts. In this letter, we offer a new CNN-based text steganalysis model that we've given the term LS-CNN. This model is able to automatically learn feature representations and capture complicated statistical correlations. The steganalysis scheme that was presented has a structure that is relatively distinct from the ones that are used in picture steganalysis. To preprocess the data, rather of employing high-pass filters, we utilize a word embedding layer to map the words into dense vectors. This allows us to get more accurate representations of the words and to extract the semantic and syntactic aspects that are associated with them. We use a number of rectangular kernels of varying sizes in the convolutional layer to extract sentence features. This is done because the semantics of language are expressed through various granularities of semantic units, such as words, phrases, and sentences. In order to better capture the features of a short text, we do this. We are able to get rich feature information to recognize short texts because the structure of the proposed model takes into consideration not only the statistical features, but also the intricate connections of words and phrases. This allows us to detect short messages. In order to further improve the performance of the suggested technique, we develop a decision strategy to be used with lengthy texts. The findings of the experiments demonstrate that the text steganalysis model that was presented is able to efficiently identify the stego text despite the length of the stego text, and achieves higher performance in terms of detection accuracy when compared to approaches that were used in the past.

The following is the structure of this letter: In Section II, a high-level introduction to steganography is presented. In Section III, our steganographic model and our decision-making procedure are presented. In Section IV, we present both the conditions of the experiment and its outcomes. In Section V, we will provide our findings and conclusions.

## 2. LITERATURE SURVEY

Text steganography can primarily be broken down into two categories: the modification-based methods, which embed secret information by modifying the cover texts and include things like synonym substitution [10], [11], syntactic transformation [12], etc., and the generation-based methods, which generate stego texts directly on the mimicking technology and include things like Markov chain-based methods [13], [14], and deep learning-based methods [15]. Both of these categories are subdivided further into subcategories

The majority of researchers are concentrating their efforts on the synonym substitution steganographic approach for modification-based steganography. This method substitutes words with their respective counterparts in order to implant messages. Taskiran et al. [1] are credited as being the first researchers to successfully identify synonym substitution steganography. They use the feature vector that is generated by a 3-gram language model and use a support vector machine to differentiate between stego texts. Chen et al. [2] make advantage of the context clustering in order to derive statistical aspects of context fitness values based on the mismatch between the replacement word and its respective contexts.

Xiang et al. [3] use relative frequency of attribute pair to generate the feature vector. This is done on the basis of the fact that the number of high frequency words is always decreased after embedding. The attribute pair consists of the synonym location and the number of its synonyms. In order to identify generation-based steganography, Yang and Cao [4] use meta characteristics (such as word length, space rate, and word frequency) in conjunction with an immune mechanism to choose the appropriate features. These meta features include: Chen et al. [5] offer a steganalysis technique called the natural frequency zone word distribution analysis (NFZ-WDA) to describe the structure of the language based on the distribution of words in the various natural frequency zones.

Although there have been numerous potential steganalysis techniques developed, the vast majority of them are tailored for a specific sort of steganographic technology. In addition to this, they depend on

specialized knowledge of the topic, such as a synonym dictionary and word frequency. In addition, the steganography techniques described above are predicated on statistical features that must be derived from a large corpus in order to function properly. As a direct consequence of this, the performance of these strategies is subpar when applied to short texts. This is what drives us to develop a global CNN-based steganography model that is capable of detecting texts of varying lengths and adapting well to the many forms of text steganography.

## 3.PROPOSED SYSTEM

In this part, the general architecture of the proposed text steganography model is described for the first time. This model can identify both long and short stego texts, therefore it can be used to analyze both types of stego texts. After that, we continue to elaborate on the CNN's structure and go through some of the design considerations that went into creating it. Finally, we show the decision approach that the text steganalysis model use in order to identify lengthy stego texts.

The general architecture of our text steganography model is seen in Figure 1. When dealing with brief texts, we first pre-process them by segmenting the words, capitalizing the first letter of each word, generating a dictionary based on the data from the training set, and encoding the words into indexes according to the order in which the words appear in the dictionary. After that, the actual index sequences are sent to the CNN (which was explained in Part B) so that it may learn the feature representations, and the predicted labels can then be acquired directly. We tokenize long texts into their sentence components that have a relatively consistent length before the data preprocessing, and we process each sentence individually. This is done for long texts because the wide range of length variation that long texts (paragraphs, chapters, books) exhibit is not conducive to the training of CNN. For this reason, we tokenize long texts. During the training phase, the CNN is trained using the aforementioned sequences together with the labels that correspond to them. During the testing step, a collection of labels of a lengthy text are predicted using the trained CNN. After that, a decision technique, which is detailed in Part C, is used to reach the ultimate determination.

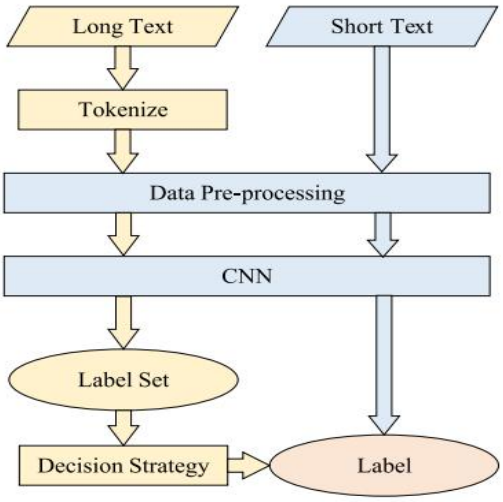


Fig. 1. Proposed text steganalysis framework.

## 4. RESULTS

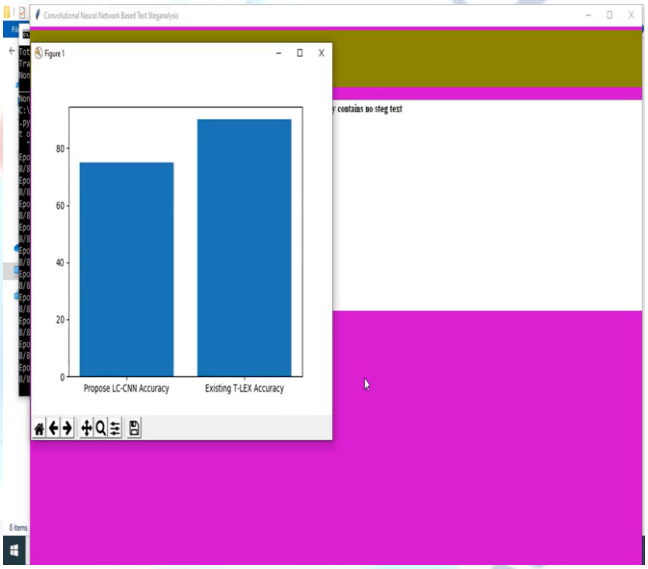


Figure 1: Comparison of accuracy with proposed and existing algorithm

## 5. CONCLUSION

In this paper, we offer a unique universal text steganography model that is based on the CNN framework. This model can train feature representations automatically and provides a superior detection performance for short texts when compared with standard statistical feature-based steganography. In contrast to the CNN that was used for picture steganography, a word embedding layer was used to extract the semantic and syntactic aspects of individual words. The sentence attributes were learnt via the use of a convolutional layer that had rectangular kernels of varying sizes.

**Conflict of interest statement**

Authors declare that they do not have any conflict of interest.

## REFERENCES


[1] C. M. Taskiran, M. Topkara, and E. J. Delp, "Attacks on lexical natural language steganography systems," in Proc. SPIE Int. Soc. Opt. Eng., 2006, vol. 6072, pp. 607209-1–607209-9.

[2] Z. Chen, L. Huang, H. Miao,W. Yang, and P.Meng, "Steganalysis against substitution-based linguistic steganography based on context clusters," Comput. Elect. Eng., vol. 37, no. 6, pp. 1071–1081, Nov. 2011.

[3] L. Xiang, X. Sun, G. Luo, and B. Xia, "Linguistic steganalysis using the features derived from synonym frequency," Multimedia Tools Appl., vol. 71, no. 3, pp. 1893–1911, Aug. 2014.

[4] H. Yang and X. Cao, "Linguistic steganalysis based on meta features and immune mechanism," Chin. J. Electron., vol. 19, no. 4, pp. 661–666, Oct. 2010.

[5] Z. Chen, L. Huang, P. Meng,W. Yang, and H. Miao, "Blind linguistic steganalysis against translation based steganography," in Proc. Int.Workshop Digit. Watermarking, Oct. 2010, pp. 251–265

. [6] P. Meng, L. Hang, Z. Chen, Y. Hu, and W. Yang, "STBS: A statistical algorithm for steganalysis of translation-based steganography," in Proc. Int. Workshop Inf. Hiding, 2010, pp. 208–220.

[7] J. Ye, J. Ni, and Y. Yi, "Deep learning hierarchical representations for image steganalysis," IEEE Trans. Inf. Forensics Secur., vol. 12, no. 11, pp. 2545–2557, Nov. 2017.

[8] G. Xu, H. Wu, and Y. Shi, "Structural design of convolutional neural networks for steganalysis," IEEE Signal Process. Lett., vol. 23, no. 5, pp. 708–712, May 2016.

[9] Y. Qian, J. Dong, W. Wang, and T. Tan, "Deep learning for steganalysis via convolutional neural networks," in Proc. SPIE Media Watermarking, Secur. Forensics, 2015, Art. no. 94090J.

[10] B. Srinivasa Rao D. Vijaya Kumar and K. Kiran Kumar, "Power quality improvement using Cuckoo search based multilevel facts controller", Journal of Engg. Research Vol.10 No. (4A) pp. 252-261, DOI: 10.36909/jer.10895.

[11] L. Huo and Y. Xiao, "Synonym substitution-based steganographic algorithm with vector distance of two-gram dependency collocations," in Proc. 2nd IEEE Int. Conf. Comput. Commun., 2016, pp. 2776–2780.