Insights into Artificial Neural Network techniques, and its Application in Steganography

Gurunath R¹ and Debabrata Samanta²

¹²Department of Computer Science, CHRIST University, Bangalore, India
E-mail: debabrata.samanta369@gmail.com

Abstract. Deep Steganography is a data concealment technology that uses artificial intelligence (AI) to automate the process of hiding and extracting information through layers of training. It enables for the automated generation of a cover depending on the concealed message. Previously, the technique depended on the existing cover to hide data, which limited the number of Steganographic characteristics available. Artificial intelligence and deep learning techniques have been used to steganography recently and the results are satisfactory. Although neural networks have demonstrated their ability to imitate human talents, it is still too early to draw comparisons between people and them. To improve their capabilities, neural networks are being employed in a number of disciplines, including steganography. Recurrent Neural Networks (RNN) is a widely used technology that automatically creates Stego-text regardless of payload volume. The features are extracted using a convolution neural network (CNN) based on the image. Perceptron, Multi-Layer Perceptron (MLP), Feed Forward Neural Network, Long Short Term Memory (LSTM) networks, and others are examples of this. In this research, we looked at all of the neural network approaches for Steganographic purposes in depth. This article also discusses the problems that each technology faces, as well as potential solutions.

Keywords: Recurrent Neural network, Convolution Neural Networks, Feed forward, LSTM, Optimization Algorithms, Deep Steganography.

1. Introduction

Steganography denotes hidden communication. The message is hidden within a cover object. Steganography is defined as a way of hiding communications using a manner that no one is aware of. Conventional Cryptography and Steganography are combined in Steganographic techniques, in which the transmitter encodes the secret message before to the whole process of communication, making it much harder for an adversary to identify hidden cipher text in a cover letter. It has been utilized by people, espionage, rulers, corporations, and militaries throughout history. Steganography is also called data hiding is the subject of several tales [1]. Data camouflage research has exploded in the last two decades, with business motivations pushing the area. A message is hidden inside a cover using data hiding techniques. The techniques’ strengths and shortcomings are characterized by a variety of characteristics such as capacity, Robustness and undetectability. Capacity is the actual number of bits concealed and correctly retrieved by the Stego system is referred to as capacity in data concealing. Robustness is, the capability of the hidden message to stay untouched when the Stego-system is transformed, such as by filtering, distortion insertion, resizing, rotation, and compression of an object. Another feature of steganography is undetectability, which means that the Stego item does not raise
any suspicion in the eyes of other parties [2]. Steganography may be used on a wide range of cover items, including text, image, audio, video, and the network. Each of these coverings serves as a concealed bearer of a secret text.

2. Types of Neural networks

Under the umbrella of Artificial Neural Networks, an explosion of technologies has occurred in the previous 10 years of Artificial Intelligence progress. The technologies of machine learning and deep learning are the most relevant in this regard. The next part explores how different neural network approaches produce quick and exact results when applied to various domains; however, in this study, we will focus on AI on Steganography. The following are the different types of neural networks, with a few of them being described in this section. They are: Perceptron, Multi Layer Perceptron, Feed forward Neural Network, Convolutional Neural Networks, Recurrent Neural Networks, Radial basis function Neural network, LSTM, Sequence to Sequence model, and Modular Neural networks [3] [4].

2.1. New Era in Neural Networks - Perceptron

In 1957, Frank Rosenblatt of Cornell Aeronautical Laboratory created Perceptron. He was an American psychologist who was impressed by our brain's neuron and its ability to learn. His Perceptron model has several inputs, a processor, and just one output. The idea was to construct a machine that could act as a biological neuron, which resulted in a software model that was then evaluated on an IBM computer before being employed for image recognition. Because the model was built for sequential purpose, it could only deal with linear data points. This resulted in complications with technology usage. However, we now use the phrase 'activation functions' to refer to the problem that existed at the time [5]. The Perceptron (Error! Reference source not found.) receives weights and a bias as inputs, which are subsequently multiplied by the weights. The products, as well as the bias, are then combined. The activation function is responsible for the final output.

For example, suppose we have three inputs $x_1, x_2, x_3$ and three corresponding weights $w_1, w_2, w_3$. The weighted sum function $f = x_1w_1 + x_2w_2 + x_3w_3$ is thus written. If the range is supplied, for example, 0 to n, we may utilize the activation function, which is expressed as:

$$\sum_{j=0}^{n} x_n w_n$$

The corresponding Figure shown below:

![Perceptron with Weighted sum function and Activation Function](image)

Figure 1 Perceptron with Weighted sum function and Activation Function

There are various types of activation functions, such as Hyperbolic Tangent and Logistic Function. The first is used to output a number between -1 and 1, whereas the second is used to output a number between 0 and 1.

The logistic and hyperbolic tangent function formula shown under:
\[ f(x) = \frac{1}{1+e^{-x}} \]  
(2)
\[ f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1 \]  
(3)
, and the typical graph for both is shown in the Figure 22 below:

Figure 2 Typical Graph of Sigmoid and tanh Activation Function curves (courtesy: Quora)

The perceptron's goal is to produce numbers that are between 0 and 1 in this case, as illustrated in the graph. Our perceptron's bias is a threshold, a number that must be reached before the output is created. As a result, the ultimate equation is,

\[ y = \sum(\text{weight} \times \text{input}) + \text{bias} \]  
(4)

Binary classifiers are used in the case of Perceptron in neural networks as supervised learning. For example, information retrieval, steganography, medical diagnosis, quality control in any organisation, and so on.

3. Multi-Layer Perceptron (MLP) Neural Network

In a continuation of the Perceptron, which has a single node and accepts weighted inputs, bias produces output depending on a rule. However, for several decades after the original invention of neural networks, there was no progress in the field of neural networks, and as a result of its incapacity, technology came to a halt. In the 1980s, processing power of computers was boosted, and new communication methods were established. Rumelhart published the MLP (Multi-layer Perceptron) method, a novel approach in neural networks, in 1986.

Figure 3 Multi-layer Perceptron Networks
4. Multi-level Feed Forward Neural (MLF) Networks
Feed forward neural networks, among other ways, are one of the most effective learning algorithms. Deep networks and multilevel perception are sometimes used interchangeably. It's a artificial neural network through which data passes for processing, such as filtering undesired data, Identifying characteristic features, and producing an output at the end [7][8].
Input layer, output layer, and hidden layer are all followed in the MLF neural networks. Source nodes are present in the input layer. This layer captures the classification of feature patterns. The dimension of the feature vector in the input determines the number of nodes. Between the input and output layers is a hidden layer. Depending on the needed performance, many hidden layers with a particular number of nodes (neurons) might be present. The hidden nodes are essential for performing sophisticated computations. Finally, the processed data is sent to the output layer. The output layer's results are directly proportional to the neural network's total performance [9][10].

5. Convolution Neural Networks
Yann LeCun named his convolution networks ConvNets when they were originally constructed in 1980. It was created to recognize handwritten digits, but it was subsequently utilized in banking and physical postal services to read the area code written on letters and cheque [11][12] [13]. It later struggled as a result of its inability to manage large amounts of data and image processing. With the introduction of deep learning technologies, particularly multi-layered neural networks, CNN may eventually be able to perform better.

6. Pseudo code
Initialize the weights with small random values
Repeat
Repeat
Select desired pattern from the training set
Fed the pattern to Input layer
Compute predicted output of the network
Compare predicted output with target value
If (predicted value <> target value)
Then update the weights
Else
Accept the value
Until all patterns selected from the training set
Until total error< criterion

7. Analysis of Neural Network Steganography
Conventional systems are slipping away owing to their limits, while neural networks represent the future of AI. Although neural nets have shown that they can mimic human skills, it is still too early to compare them to humans. The neural networks are being used in a variety of fields to increase their capabilities. In terms of improving payload capacity, undetectability, and robustness, steganography can benefit from neural networks. The study of Steganography papers count collected from Google Scholar from 2016 to 2020, a total of 5 years of data, is presented in this section. The focus of the analysis is on neural network techniques used in text and image Steganography [14][15][16].

With the search phrase "Steganography and Neural Networks," we looked at the first 20 publications of each year starting in 2016. Only publications relating to text and image steganography were included in the 20-article collection. Other forms of Steganography, such as audio, video, and network, have not been studied. In the following years, 2017, 2018, 2019, and 2020, we did the same thing. We analyzed a total of 100 articles, focusing solely on the various Neural Network methods employed on text and images [17][18][19].
The following stacked graph depicts the graphical representation of the prior tabular data. For the past 5 years, researchers have focused their efforts on Convolutional Neural Steganography, followed by Generative Adversarial Networks (which are extensively used in image, video, and voice creation), recurrent neural networks (Text Generation), and deep neural networks. Although the others have fewer publications, they have more research possibilities.

**Figure 4** Articles on Steganography published during last five years on based on Neural Network Technologies.

For last five years data shows, researchers are trying to explore a number of neural network methods into their domain. In Steganography as per the analyzed data CNN has gained superiority (Figure 4), because traditional classification methods impact the pixel characteristics between them when image processing; CNN is a prominent image classification approach, deals this issue by employing down sampling.

The second in the graph, GAN stands for "generative adversarial network." It is a Machine Learning framework. It is a well-known method that employs two networks: a generative and a discriminator network. The former network continually generates pictures with certain properties, whilst the latter attempts to distinguish correct and wrong images. This method creates pictures that are identical to real images.

RNN and Deep Neural networks are also in the race, with more study being conducted in these areas of neural networks based on steganography. Even if there are fewer approaches, they have a lot of research potential.
We also looked at the number of articles published by each form of Steganography over the same years, as shown in the table and pie chart below (Figure 6). Image Steganography has once again claimed the lead in the number of articles with 65. The text Steganography, on the other hand, contains a relatively small number of articles, just 7. The argument is that concealing data in text is more difficult than hiding data in an image.

Because image data is a binary file containing a large number of repetition bits. Furthermore, the LSB replacement embedding approach in images is relatively easy and has a higher payload capacity. Text concealing, on the other hand, is always difficult owing to the fewer repeated bits. As a result, the amount of concealment is fairly limited. As a result, the number of papers published during the last five years is less.

8. Conclusion
The term "deep steganography" refers to steganography techniques that employ Neural Network technology. What is the aim of deep steganography? The answer may be found in the amount of publications that have occurred in recent years, as well as the difficulties with traditional data-hiding techniques. Deep Steganography popularity is on the rise, and it's producing incredible results. The article's main goal is to look at how Steganography has progressed in relation to neural networks. These techniques assist in the creation of the cover carrier depending on the message size, providing high payload capacity, robustness, and the absence of steganalysis. In this work, several ANN methods are described in depth. RNN is a type of neural network that is used to train a model in natural
language processing to generate a series of words. Forward and backward networks can be used in RNN. Long-term dependencies are a difficulty for RNNs; it’s difficult to learn to link to information when there’s a lot of context. To solve this RNN issue, the LSTM was designed. LSTM, on the other hand, takes longer time to train the model and has an over-fitting problem. The updating of weights and learning rate is excessively sensitive; any minor change causes LSTM to provide incorrect results. CNN is a promising image classification technique, but it tends to produce incorrect results when there is a lack of training data. CNN is unable to process textual data. To reduce loss, neural networks employ optimization techniques. Gradient descent works best with small data sets because reaching local minima takes too long otherwise.

References

[1] A. A., B. B. Zaidan, and Anas Majeed. Zaidan, "High securing cover-file of hidden data using statistical technique and AES encryption algorithm,” World Academy of Science Engineering and Technology (WASET), pp. 468-479., 2009.

[2] A. W., Hameed, S. A., Zaidan, B. B., Al-Khateeb, W. F., Khalifa, O. O., Zaidan, A. A., & Gunawan, T. S. Naji, "Novel framework for hidden data in the image page within executable file using computation between advanced encryption standard and distortion techniques.,” arXiv preprint arXiv:0908.0216., 2009.

[3] M., Besacier, L., & Verbeek, J. Elbayad, "Pervasive attention: 2d convolutional neural networks for sequence-to-sequence prediction," arXiv preprint arXiv:1808.03867, 2018.

[4] C., Miaofen, Z., Honghao, Y., & Yan, L. Guojin, "Application of neural networks in image definition recognition.,” in IEEE International Conference on Signal Processing and Communications, 2007, pp. 1207-1210.

[5] F., Macchiavello, C., Gerace, D., & Bajoni, D. Tacchino, "An artificial neuron implemented on an actual quantum processor.,” npj Quantum Information, vol. 5, no. 1, pp. 1-8, 2019. [Online]. https://towardsdatascience.com/what-is-a-perceptron-basics-of-neural-networks-c4cfea20c590

[6] P. M., & Tatnall, A. R. Atkinson, "Introduction neural networks in remote sensing,” International Journal of remote sensing, vol. 18, no. 4, pp. 699-709, 1997.

[7] G. M., Foody, "Land cover classification using an artificial neural network with ancillary,” International Journal of Geographical Information Systems, vol. 9, 1995.

[8] Gurunath R, Debabrata Samanta, Studies on Encrypted Secret Data Storage Techniques Analogous to Steganography, International Journal of Advanced Science and Technology, Vol. 29, No.2, (2020), pp.3705-3711.

[9] D Sarkhel, D Deka, Debabrata Samanta , Kumudavalli M.V and Dac-Nhuong Le, GUI based Percentage Analysis for Cure Breast Cancer Survivors, Proc. of The 7th International Conference On Frontiers Of Intelligent Computing: Theory And Application (FICTA 2018) , Nov 29 - 30, 2018 at Duy Tan University, Da Nang, Viet Nam.@ Springer AISC

[10] R Gurunath , M Agarwal, A Nandi, Debabrata Samanta , An Overview: Security Issue in IoT Network, Proc. of IEEE– 2nd International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2018), 30-31 August 2018 by SCAD Institute of Technology at Coimbatore,India.

[11] M Agarwal, D Sarkhel, Debabrata Samanta, Architecture of ‘Loop’ without ‘for’ & ‘while’, Proc. of IEEE – 2nd International Conference on Inventive Systems and Control (ICISC 2018) @IEEE,19-20 January 2018 at JCT college of Engineering and Technology, Coimbatore, India.

[12] Dhanush V, Mahendra A R, Kumudavalli MV,Debabrata Samanta , Application of deep learning technique for automatic data exchange with Air-Gapped Systems and its Security Concerns, Proc. of IEEE International Conference on Computing Methodologies and Communication [ICCMC 2017] @IEEE , 18-19, July 2017 , Erode.

[13] R Kumar, Rishabh K,Debabrata Samanta , M Paul, CM Vijaya Kumar, A Combining approach using DFT and FIR filter to enhance Impulse response, Proc. of IEEE International Conference on Computing Methodologies and Communication [ICCMC 2017] @IEEE , 18-19,
July 2017, Erode.

[14] G Ghosh, Debabrata Samanta, M Paul, N Kumar Janghel, Hiding Based Message Communication techniques depends on Divide and Conquer Approach, Proc. of IEEE International Conference on Computing Methodologies and Communication [ICCMC 2017] @IEEE, 18-19, July 2017, Erode.

[15] R. K. Singh, T. Begum, L. Borah, Debabrata Samanta, Text Encryption: Character Jumbling, Proc. of IEEE International Conference on Inventive Systems and Control [ICISC 2017] @IEEE, 19-20 January 2017, Coimbatore. 978-1-5090-4715-4/17/$31.00 @ 2017 IEEE.

[16] R. P., Lippmann, "An introduction to computing with neural nets," I.E.E.E. ASSP, vol. 2, 1987.

[17] S. and Tatnall, A. R. L. Co* te’, "The Hop®eld neural network as a tool for feature," International Journal of Remote Sensing, vol. 18, 1997.

[18] W. J., & Chen, F. W Blackwell, "Neural networks in atmospheric remote sensing," Artech House, 2009.