iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** IV **Month of publication:** April 2021

DOI: https://doi.org/10.22214/ijraset.2021.33459

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue IV Apr 2021- Available at www.ijraset.com

Steganalysis and Encryption Detection using Deep Learning

Pampana Venkata Srimukh[1], Aditya Kumar Sinha[2], Saket Kumar Jha[3]
[1]*Mtech CSE (Artificial Intelligence and Machine Learning), Institute of Technology University, Chennai*
[2]*Director and Centre Head, C-DAC Patna*
[3]*Project Engineer C-DAC Patna*

Abstract: Data is the building block of technological marvels in 21st century. It has become a precious gem and is probably considered invaluable than any other. Today's crimes involve mostly on data theft, which has become a thorn on the side not only for an individual using a computer but also large corporates and governments. Out of all the crimes in relation to pilfering the data, Steganography stands tall. The recent advancements in the Steganography, The possibility of identifying a payload in a multimedia platform has become complicated. With the help of Deep Learning methods, A State-of-the-Art steganalysis model needs to be developed to rival this dynamic threat.
Keywords: Steganography, steganalysis, deep learning, cyber forensics

I. INTRODUCTION

Steganography means the art of concealing a piece of information in an object. It had been in existence that dates back to 440 BC. The word Steganography is derived from greek words *steganos* meaning 'covered or concealed' and *graphia* meaning 'writing'. The Roman Empire used steganography extensively. They used slave's head by shaving off the hair and etched the information on his scalp. Later when the hair has grown back, the slave was sent to deliver the message for a person who in turn removes the hair and read the message. Although the concept existed since ancient times, it was in 1499 the term 'Steganography' coined by Johannes Trithemius in his book 'Steganographia'. In WWII, microdots were used to send secret messages to espionage departments. In the 21st century steganography is mostly used in compromising sensitive information belonging to corporate sector in stealing technology to use it for illegal purposes. The art of deciphering the payload(secret message in an object) is called Steganalysis. It is a technique employed in retrieving the hidden messages across the multimedia platform in today's world. This technique is much more complex than the cryptanalysis. In cryptanalysis, a person was given a task of retrieving the information from an impeding message. While in steganalysis, the person needs to find out the type of file or object from a pile of evidence collected and try to reconstruct the secret message hidden inside. In older days, Steganalysis is performed from art. Now since the various sources of multimedia platform available in the palm of user, this was becoming an ever evolving problem. Steganalysis can be performed on many things but the most common were performed on images, audio clips and videos. Due to the availability of images in plethora kudos to internet, finding the cover image would can be flexible. Least significant bit aka LSB technique is the key element in applying steganography. It is the simplest method to embed secret information. It uses the least significant bits in pixels of the image and switch them with the information which is converted to binary form. The LSB's of an image might be the left most end or at right most end of the binary digit. Another form of embedding information is by using Discrete Cosine Transform(DCT). It is used for JPEG compression. JPEG is a lossy compression technique that reduces the size of the digital images. DCT will divide different parts of an image based on their significance. DCT metamorphoses an image which is in spatial domain to frequency domain.

II. SURVEY

Human intervention in situations that decides the fate of perpetrator of certain crimes, This proposed DLCF framework[1] would be a great tool in reducing the human effort. This is a tool that can sift through large amount of data in no time to identify the stolen data. It also suggests the transition of cyber forensics domain that in future this could wide open the research and application of DL in this domain. Although the utilization of machine learning techniques in cyber forensics have already proven that the efficiency of the machines can be improved widely. ML techniques as suggested by prerak bhatt[2] is a living proof. The residual network employment has yielded optimum results in finding the missing documents that were stolen. The collaboration of deep neural networks and reinforcement learning[8] in training the datasets to extract the most possible regions that are targeted to be steganographic epicenter has taken a new step in the growth of this application in deep learning. Creation of multi agents[3] to coordinate the tasks were set for different agents using JADE. MADIK is helpful in giving the suggestions to the specialists when a recommendation is needed. DeepSign[4] is an auto detecting system of malware signature based on deep belief network(DBN).

The authors have used cuckoo sandbox to collect the malwares and then later on converted them to binary files which then it was trained on DBN. A review[5] on deep learning based Steganalysis and its evolution from 2015 to present was performed.

IoT is another domain where the security measures to protect the data needs to be researched. The authors[6] proposed a hybrid model to alleviate the cyber threats. An Intrusion Detection System(IDS) based on Convolution Neural Network was the basic idea that will classify the threats in IoT devices. The provision of good security measures in IoT devices is seeming to be far fetched. An intricate detection system[7] based on tensorflow, scikit-Learn and seaborn was developed that will detect the Denial of Service(DoS) attacks that use IoT devices as a medium to launch full pledged strike on a website. This system will check whether there is any indication of the cyber attack on the device and then it will be granted access to the database. [9] Quantization technique was implemented to compress the memory of the images. This is compared with the non-compressed image dataset to check if the neural networks can give the same accuracy. The experiment was performed in two cases: once with DCT and without DCT. Even though the difference of accuracy is less, it gave a conclusion that compression technique is vital to reduce the effort on the machine and can yield the same result. In this paper[10], the author evaluated different neural networks to StegDetect for image classification. StegDetect detects the stego images. But in terms of accurately classifying the images, neural networks especially SE-ResNets were much better classifiers than the StegDetect.

III. REVIEW SUMMARY

The DLCF framework promises to be a vital tool that can ease the loss of time factor, which is one of the challenge that needs to be addressed although it is a proposed framework, the concept of incorporating deep learning may very well pave the way to create good techniques. Machine Learning also promises on developing a good ANN based machine to but deep learning outshines the traditional machine learning techniques in terms of performance and complexities although it was able to identify data that was present in the system but it does not have the computational power to perform on a large data source. MADIK is a powerful intelligent tool kit that can analyze the data via its intelligent systems, specifically designed that each system will perform the task of identification coordinately. This kit might be a powerful tool since it is completely based on unsupervised learning. The application of unsupervised technique in deep learning such as incorporating deep belief network in identifying the malware on its own, This system has proven that deep learning models especially DBN is able to achieve high accuracy rate in identifying the malware. Steganalysis is an old trick in the book but very effective in concealing the data within the digital media platform, with the literature survey stating the evolution of research by employing deep learning techniques suggest that although there is a progress in this area but much to be researched to overcome one of the biggest challenges of cyberforensics. Security in IoT based devices are poor when compared to the normal computers but in the case of applying ML techniques in predicting the attacks using previous history and pattern of attacks is a conventional method. If this process can be tested in terms of creating multi agent system to do the work, it will be yielding greater results. The comparison of different algorithms and their performance evaluation has yielded better results and this might be useful in stretching the boundaries of deep learning further. The incorporation of deep learning and reinforcement learning for a self seeking steganalysis mechanism is a new approach in the field of steganalysis and also advanced. The concept of quantization in steganalysis models will prove beneficial for the machine to eradicate the problem of high computational power yet the performance will be satisfactory. Here the comparison test of different deep learning models with stegdetect shows that much to be done in the field of steganalysis.

IV. METHODOLOGY

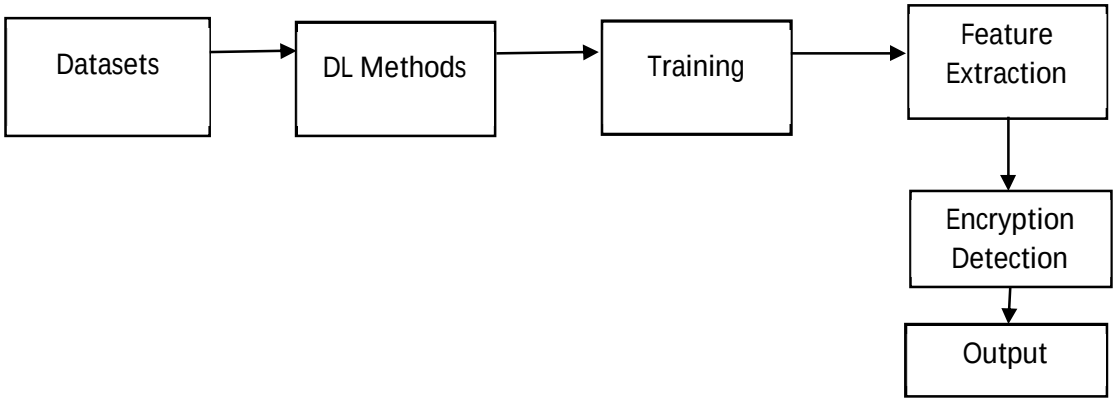


Figure: Architecture of the proposed methodology

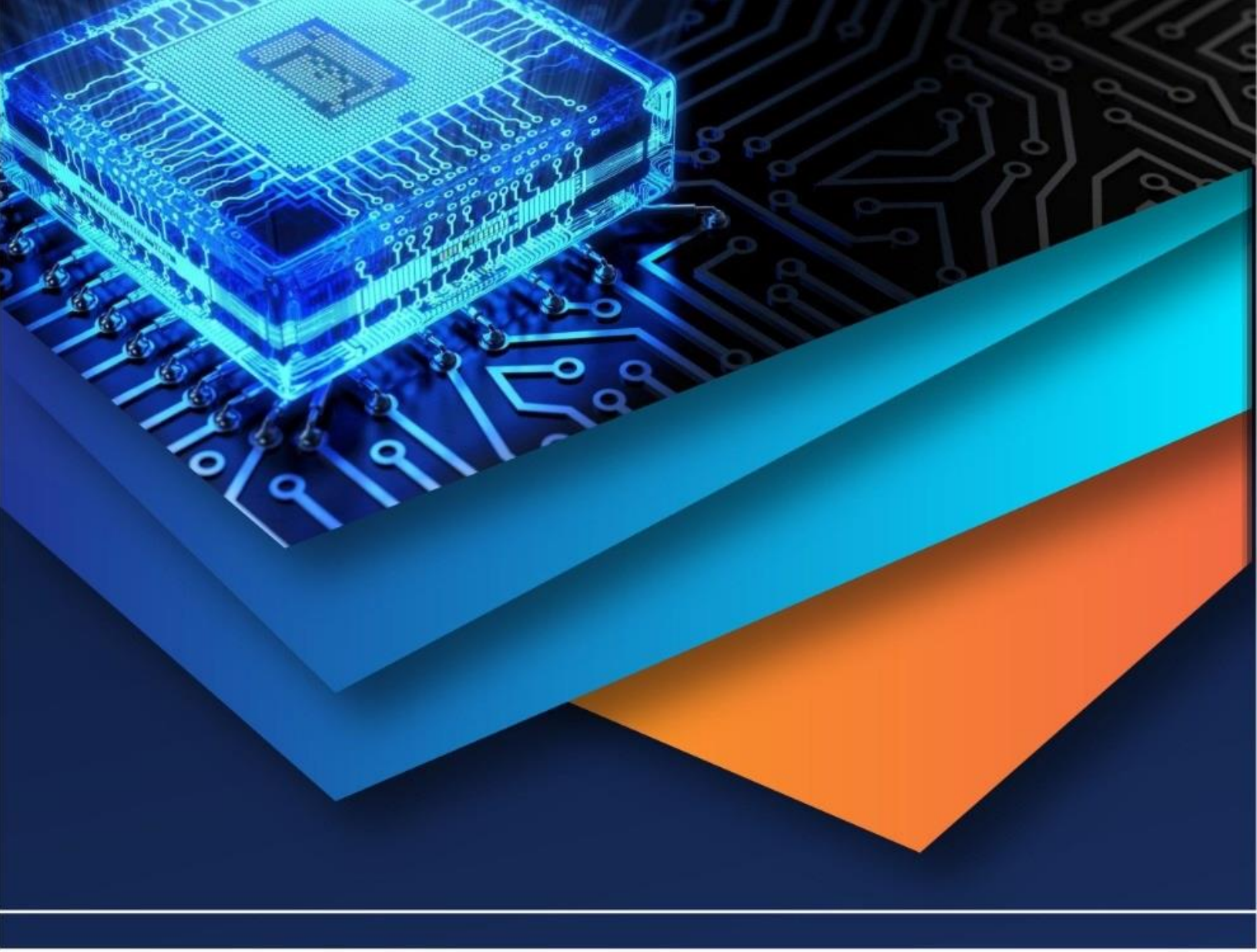
The proposition of this methodology is to identify if there's any payload inside an object. The existing models are trying to work with the challenges posed in detecting the payload that was embedded using the DCT coefficients. This method also takes an insight that what if the payload is encrypted even before it is being etched into the object was already encrypted just in case to protect the payload even if it was detected?. By taking this as a base to our system, we tend to experiment with images as our carrier of the payload. The architecture proposed here states that, a dataset comprises of both the original and stego images are needed to be trained with DL methods. By training, the features are going to be extracted and then it passes through an encryption detection phase where it would trace out whether the payload is already encrypted before being embedded. If it detects that the message was in encryption then it will be deciphered and then produce the output. Here we are looking into alaska2 dataset. It is specifically built to train and test the steganographic images. Convolutional Neural Networks(CNN) is the primary method that were being considered in this methodology to identify any differences between the Cover image and the payload image. Although, a deep belief network which works on binary latent variables would be much interesting to experiment on. An Encryption detection is a system which detects whether the payload is encrypted even before it is encoded in the image.

V. CONCLUSION

Steganalysis is going to play a vital role in recovering the stolen data. Due to the advanced steganography tools which threatens the protection, Deep Learning employed Steganalysis might provide the solution to build a barrier to counter the felonies committed by the cyber criminals. We hope the proposed system will provide the solution to detect the DCT based steganography and makes it easier to reduce the time and efforts before being compromised.

REFERENCES

[1] Karie, Nickson & Kebande, Victor & Venter, H.. (2019). Diverging deep learning cognitive computing techniques into cyber forensics. Forensic Science International: Synergy. 1.10.1016/j.fsisyn.2019.03.006.
[2] Bhatt, Prerak. (2017). MACHINE LEARNING FORENSICS:A NEW BRANCH OF DIGITAL FORENSICS. International Journal of Advanced Research in Computer Science. 8. 217-222. 10.26483/ijarcs.v8i8.4613.
[3] Hoelz, Bruno & Ralha, CÃ©lia & Geeverghese, Rajiv. (2009). Artificial intelligence applied to computer forensics. Proceedings of the ACM Symposium on Applied Computing. 883-888. 10.1145/1529282.1529471.
[4] O. E. David and N. S. Netanyahu, "DeepSign: Deep learning for automatic malware signature generation and classification," *2015 International Joint Conference on Neural Networks (IJCNN)*, Killarney, 2015, pp. 1-8, doi: 10.1109/IJCNN.2015.7280815.
[5] T. Reinel, R. Raúl and I. Gustavo, "Deep Learning Applied to Steganalysis of Digital Images: A Systematic Review," in *IEEE Access*, vol. 7, pp. 68970-68990, 2019, doi: 10.1109/ACCESS.2019.2918086.
[6] Temechu G. Zewdie and Anteneh Girma, "IOT Security and the Role of AI/ML to Combat Emerging Cyber Threats in Cloud Computing Environment," Issues in Information Systems, Volume 21, Issue 4, pp. 253-263, 2020
[7] Susilo, Bambang & Sari, Riri. (2020). Intrusion Detection in IoT Networks Using Deep Learning Algorithm. Information. 11. 279. 10.3390/info11050279
[8] D. Hu, S. Zhou, Q. Shen, S. Zheng, Z. Zhao and Y. Fan, "Digital Image Steganalysis Based on Visual Attention and Deep Reinforcement Learning," in IEEE Access, vol. 7, pp. 25924-25935, 2019, doi: 10.1109/ACCESS.2019.2900076.
[9] X. Wu, Z. Shao, P. Ou and S. Tan, "Application of quantisation-based deep-learning model compression in JPEG image steganalysis," in The Journal of Engineering, vol. 2018, no. 16, pp. 1402-1406, 11 2018, doi: 10.1049/joe.2018.8299.
[10] Mamada, Naoya. "Image Steganalysis with Very Deep Convolutional Neural Networks." CLEF (2019).



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY