Technological University Dublin
ARROW@TU Dublin

Conference papers | School of Electrical and Electronic Engineering

2020

Modern Techniques for Discovering Digital Steganography

Michael Hegarty

Anthony J. Keane

Follow this and additional works at: https://arrow.tudublin.ie/engscheleart

 Part of the Electrical and Computer Engineering Commons

This Conference Paper is brought to you for free and open access by the School of Electrical and Electronic Engineering at ARROW@TU Dublin. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, gerard.connolly@tudublin.ie.

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 4.0 License

Modern Techniques for Discovering Digital Steganography

Michael T. Hegarty and Anthony J. Keane
TU-Dublin, Blanchardstown, Ireland
Michael.hegarty@tudublin.ie
Anthony.keane@tudublin.ie
DOI: 10.34190/EWS.20.504

Abstract: Digital steganography can be difficult to detect and as such is an ideal way of engaging in covert communications across the Internet. This research paper is a work-in-progress report on instances of steganography that were identified on websites on the Internet including some from the DarkWeb using the application of new methods of deep learning algorithms. This approach to the identification of Least Significant Bit (LSB) Steganography using Convolutional Neural Networks (CNN) has demonstrated some efficiency for image classification. The CNN algorithm was trained using datasets of images with known steganography and then applied to datasets with images to identify concealed data. The algorithm was trained using 5000 clean images and 5000 Steganography images. With the correct configurations made to the deep learning algorithms, positive results were obtained demonstrating a greater speed, accuracy and fewer false positives than the current steganalysis tools.

Keywords: Steganography, Deep Learning, JPEG, LSB, Darknet, Openpuff

1. Introduction

Steganography can be explained as the art and science of writing hidden messages. Using this method no one apart from the sender and intended recipient suspects the existence of the message. In the modern digital world, Steganography methods are many as information can be hidden not only within graph images but also within other digital carriers such as video and audio (Kessler, 2015).

According to (Zielińska *et al,* 2014), technology makes it easy to perform digitalized methods of concealing information with many of the applications available for free from internet portals. Currently, there are approximately 1200 tools available to create different forms of steganography (Hegarty, 2018). These tools can be easily used by criminals to communicate undetected and challenges the cyber security specialists to establish if covert communication has occurred. Digital evidence of communication may never be detected on a suspect's machine due to it being concealed within a carrier file (Hunt, 2012). Richer (2015) suggests that Steganography can be used to communicate with complete freedom under conditions that are monitored.

While there is the suspicion that terror groups are using technologies like Steganography to help conceal their communications on the Internet, there is little solid evidence to support that claim. In May 2011 a 22-year-old Austrian citizen Maqsood Lodin was stopped and questioned in Germany. He was traveling from Pakistan to Berlin via Hungary. During a search, a memory stick was found in his underpants. The memory stick contained two pornographic videos. Over 100 documents were concealed within data on the storage device. They contained secret documents about operational details and training manuals for al-Qaeda. Its source came from a German Magazine "Die Zeit". The journalist and author of the article Yassin Musharbash has published an investigative report on his personal blog that includes a summary of the findings (Musharbash, 2012).

There are examples of the use of Steganography in industrial espionage when in 2018, Zheng Xiaoqing and Zhang Zhaoxi were both indicted for their role in stealing proprietary trade secret information related to GE's turbine technology. The indictment identified highly sophisticated steganography measures deployed to hide the stolen GE trade secrets in JPeg images files and sending them to a personal Hotmail account (Lynch and Shepardson, 2019).

2. Digital Steganalysis

Steganalysis involves examining several parameters of media type for Steganography to check for hidden data. Statistical Steganalysis is when alterations are applied to a carrier file, there is a change in the statistical information such as changes made to the LSB of the original image file in order to hide information (Ker *et al*., 2013).

The detection of files with hidden data required that the servers/applications that store the file do not change the files or the secret payload can be destroyed. Ebay does this to files to fit its requirements and renders the files useless for retrieving hidden information. Web-portals on the clear and darknet were identified as places that steganography could survive as the files are stored in original formats. Some of these web-portals that were researched are:

- Draugiem: The main social media platform in Latvia with 2.6 million users
- VKontakte (VK): Russia's largest social media platform claiming to have 500 million registered accounts
- ISIS Darknet: wpcxzq4ykmsxpacm.onion a propaganda outlet for Islamic state on the darknet

A dataset of 3 million images was harvested from clear-net and dark-net webportals. This data was analysed using existing steganalysis tools such as Stegdetect, Stegexpose, Outguess, and Virtual Steganographic Laboratory (VSL). Steganalysis tools will return a probability and categorise findings. Highly probable steganography images were then further analysed. Of the 3 million images analysed 0.02% (600 images) returned as suspicious of concealing data.

3. Machine Learning for Steganalysis

According to (Giarimpampa, 2018), *"Neural Networks have proven to be a useful tool in image Steganalysis".* The use of both Support Vector Machine (SVM) and Artificial Neural Networks (ANN) algorithms have become the predominant modern Steganalysis trend.

Using Python, Anaconda, on the Spyder IDE, 5000 clean images with no concealed data was developed. 5000 steganography images were created using OpenPuff steganography tool. OpenPuff is the most downloaded steganography tool according to Softpedia with 60,352 downloads as of September 2019. Figure 1 displays an original image on the left and the carrier image (with a secret massage) on the right, this secret message was created using OpenPuff steganography software. Figure 2 presents a sample of the HEX code and the changes OpenPuff has created to hide the secret message in the LSB.

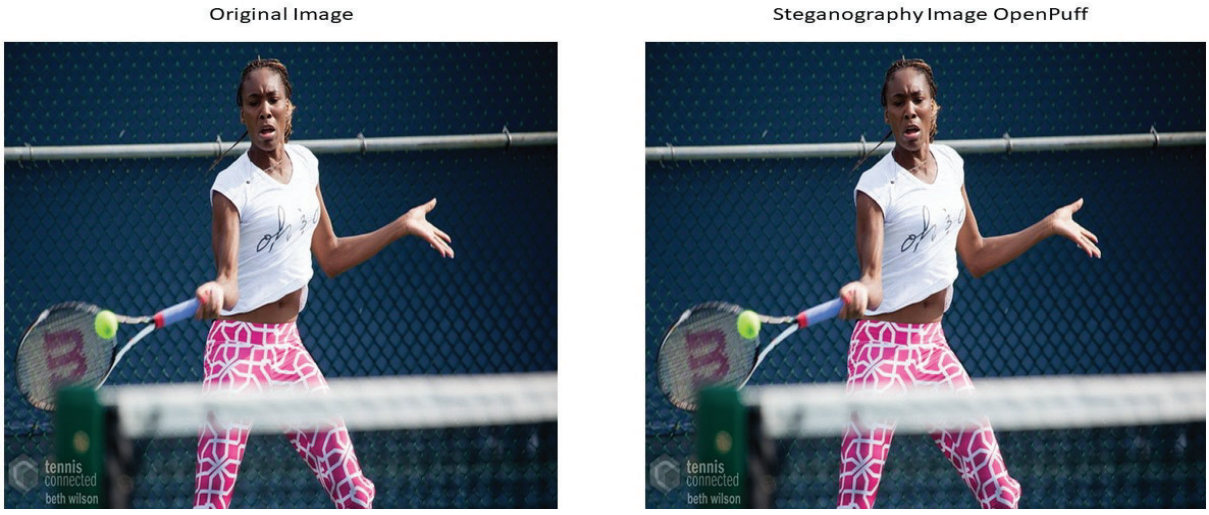

Figure 1: Sample original image and same image with hidden message

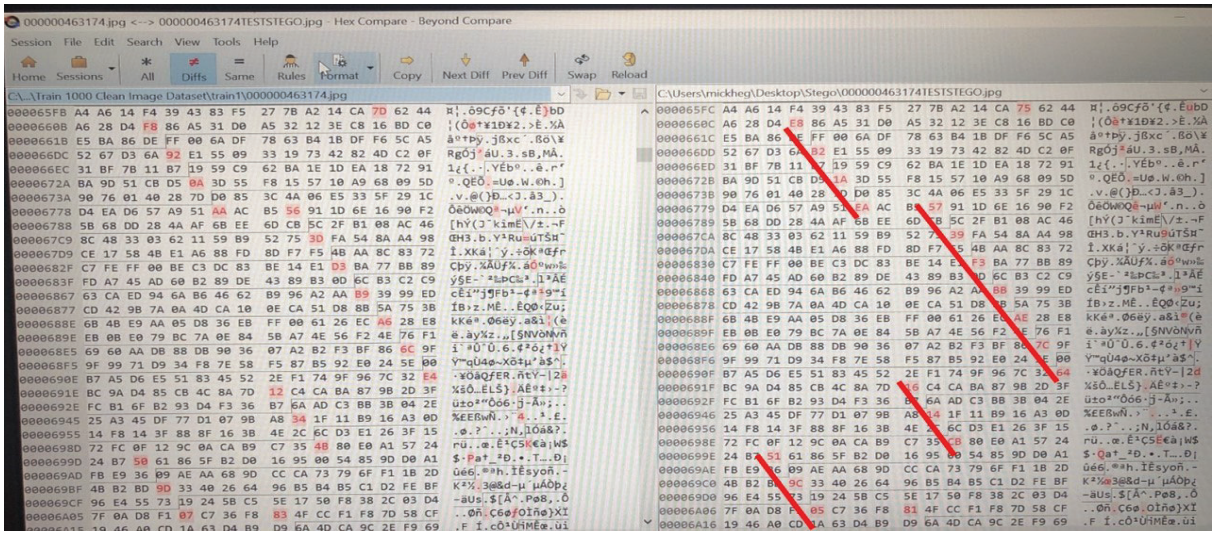

Figure 2: HEX code of original image and HEX code of image with hidden message

The Deep Learning algorithm Convolutional Neural Networks (CNN) was deployed to train and test the image datasets. The algorithm has been used for steganalysis purposes because of efficiency in image classification (Liu., Yang, and Kang, 2017).

The CNN algorithm Python code is designed to be able to perform image classification on the datasets being trained and tested. The convolution code will extract the features and learn the image feature matrix also known as the feature map. The Pooling code will reduce the dimensions of each feature map but keep the most important data. The Flattening code will convert the matrix into a linear array and Fully Connecting code will connect the convolutional network to the neural network and then compile the network. The Epoch code is set to 100 which means the dataset will pass through the neural network 100 times which will result in an increase in accuracy and a decrease in loss .

The following metrics were used:

- The **Training Accuracy** represents the accuracy that is applied to the CNN model on the training data function as a percentage and how well the CNN model was trained.
- The **Training Loss** represents the error on the training set of that data. A decrease in the Training Loss means the CNN model is improving.
- The **Validation Accuracy** is used after the CNN has been trained for adjusting the networks hyper-parameters and comparing how the changes to them affect the predictive accuracy of the CNN model.
- The **Validation Loss** is used after the CNN has been trained and represents the errors discovered after the test was used to test the predictive accuracy of the trained CNN model.

4. Initial Experiments

The CNN utilised the algorithm to train 5000 clean images and test 5000 stego-images testing different configurations to decrease the loss and increase validation

Experiments 1-3 were tested using the Final Layer Activation Function Sigmoid which is used for binary classification in logistical regression modelling.

Experiments 4-6 were tested using the Final Layer Activation Function Softmax which is used for multiple classifications in logistical regression modelling.

Table 1: Results of Experiments

Results of the Current Experiments to date								
Experiment Number	Final Layer Activation Function	1st Layer Filter	Epoch per Steps	Validation Steps	Training Loss	Validation Loss	Training Accuracy	Validation Accuracy
Experiment 1	Sigmoid	32, size of 3x3	80	8	0.011771 Decrease	5.19850 Increase	0.99558 Increase	0.57692 Increase
Experiment 2	Sigmoid	32, size of 3x3	800	80	0.00416 Decrease	6.78965 Increase	0.99908 Increase	0.56939 Increase
Experiment 3	Sigmoid	64, size of 7x7	80	8	0.00809 Decrease	5.78904 Increase	0.99759 Increase	0.57692 Increase
Experiment 4	Softmax	64, size of 7x7	80	8	8.22109 Consistent	8.78875 Consistent	0.48432 Consistent	0.44872 Consistent
Experiment 5	Softmax	32, size of 3x3	80	8	8.19546 Consistent	9.06127 Consistent	0.48593 Consistent	0.43162 Consistent
Experiment 6	Softmax	32, size of 3x3	800	80	8.09873 Consistent	8.55683 Consistent	0.49200 Consistent	0.46327 Consistent

4.1 Comparison of Experiment 1 & 3: Sigmoid

Results showed a decrease in the Training Loss of 0.003681 but an increase of 0.59054 in the Validation Loss. The Training Accuracy increased by 0.00201 and the Validation Accuracy was identical 0.57692. This comparison of results indicates when configuring the 1st Layer Filter of "32, size of 3x3" to "64, size of 7x7" it can improve the CNN model by decreasing the Training Loss and increasing the Training Accuracy. The experiment had a Validation prediction loss increase of 0.59054% but showed improvement when compared to the 1st Comparison as the Validation Accuracy remained balanced.

4.2 Comparison of Experiment 5 & 6: Softmax

Examining and comparing experiments 5 and 6, results showed a decrease in the Training Loss of 0.9673 and a decrease in the Validation Loss of 0.50444. The Training Accuracy increased by 0.00607 and the Validation Accuracy increased by 0.03165. This comparison of results indicates when configuring settings for the Epoch per steps from 80 to 800 and the Validation steps from 8 to 80 it can improve the CNN model results as both the training and validation Losses decreased. Also both the training and validation Accuracy results were increased.

4.3 Comparison of Experiment 2 & 6

Comparing experiment 2 with Sigmoid configurations and experiment 6 with Softmax configurations. Using the configurations for the Epoch per steps 800 and the Validation steps 80 with the 1st Layer Filter of "32, size of 3x3" being set. The results showed the Training Loss increased by 8.09873 from Sigmoid to Softmax and the Validation Loss increased by 1.76718. The results also showed a decrease in Training Accuracy by 0.50708 and a decrease in the Validation Accuracy by 0.10612. The results in this comparison are significantly larger than the other experiments as the Sigmoid is used for binary classification and the Softmax is used for multiple classifications.

5. Conclusions

As the configuring the Epoch per steps to 800 plus Validation steps to 80, and configuring 1st Layer Filter of "32, size of 3x3" to "64, size of 7x7" have showed that they can improve the performance of the CNN model for both Accuracy and Loss.

These low results for Softmax for the Validation Accuracy, which show a decrease of 0.1% and a Validation Loss increase of 2 on average when all 6 experiments were compared may indicate detection of steganalysis when compared to the Sigmoid results.

When implementing the deep learning CNN algorithm, results showed that increasing configurations such as the 1st Layer Filter, size, the epoch steps and the validation steps in the CNN code. The CNN model can improve the training and validation accuracy, also decreasing the loss in the training and validation of the CNN model. This leads to improved accuracy in the detection of OpenPuff steganography in a large dataset

This work in progress will now examine increased secret message size within the carrier image and various settings in the CNN model to improve accuracy.

References

Giarimpampa, D., 2018. Blind Image Steganalytic Optimization by using Machine Learning.

Types of Machine Learning. (2018). https://towardsdatascience.com/machine-learning-for-beginners-d247a9420dab [Accessed 14 Aug. 2019]

Hegarty, M. (2018), “Steganography, The World of Secret Communications”. CreateSpace Publishing Platform. ISBN-10: 1986125424

Hunt, R., 2012, December. New developments in network forensics—Tools and techniques. In 2012 18th IEEE International Conference on Networks (ICON) (pp. 376-381). IEEE.

Kessler, Gary C. "An Overview of Steganography for the Computer Forensics Examiner (Updated Version, February 2015)." (2015).

Ker, A.D., Bas, P., Böhme, R., Cogranne, R., Craver, S., Filler, T., Fridrich, J. and Pevný, T., 2013, June. Moving steganography and steganalysis from the laboratory into the real world. In Proceedings of the first ACM workshop on Information hiding and multimedia security (pp. 45-58). ACM.

Liu, K., Yang, J. and Kang, X., 2017, May. Ensemble of CNN and rich model for steganalysis. In 2017 International Conference on Systems, Signals and Image Processing (IWSSIP) (pp. 1-5). IEEE.

Lynch, S. Shepardson, D. “U.S. accuses pair of stealing secrets, spying on GE to aid China” Reuters News Agency, April 23, 2019. https://www.reuters.com/article/us-usa-justice-ge/us-accuses-pair-of-stealing-secrets-spying-on-ge-to-aid-china-idUSKCN1RZ24O [Accessed: 15th August 2019]

Morrison, B. (2004) “Ex-USA TODAY reporter faked major stories” USA Today. 3/19/2004 4:13 AM http://usatoday30.usatoday.com/news/2004-03-18-2004-03-18_kelleymain_x.htm [Accessed:10th September 2019]

Provos N and Honeyman P. 2001. Detecting Steganographic Content on the Internet: Center for Information Technology Integration University of Michigan (August 2001).

Richer, P., (Updated Version, February 2015). Steganalysis: Detecting hidden information with computer forensic analysis. SANS/GIAC Practical Assignment for GSEC Certification, SANS Institute, 6.

Van Heerden, H. "A First Course in Ethical Hacking" Second Edition, 07 November 2014

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.