International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 08, August-2021

Image Stegnalysis through Deep Learning and LSB

Nitin Pramod Ranjan
18BCE0272, School of Computer Science
Vellore Institure of Technology
Vellore,India

***Abstract*— Image Steganalysis is the detection of the message within the cover-image. It is often one of the most difficult processes in image processing because the information encoded in the image can often be fully retrieved only when the algorithm used to encode and embed it is known. This would mean that there is a need for a model that can understand the encryption or stenographic models. Most of these models rely upon deep learning alone. The author proposes the use of a RESNET-50 deep learning model with additional filters. The model proposed achieves an average accuracy of 75 percent on the datasets tested.**

Keywords*— Stegnalysis, encryption, deep learning, RESNET-50, LSB filtering

I. INTRODUCTION

Steganalysis is the study of detecting hidden messages using steganography. This is analogous to crypto-analysis and cryp-tography. Steganalysis can prevent the covert communication by analyzing the statistical distribution of the carrier, detecting and even extracting the secret messages embedded in the digital image. However, it is very difficult to accurately model the original carrier image using statistical methods, which will undoubtedly hinder the detection of secret information in the steganographic image. So, most practical applications rely on machine learning models. In 2015 Imagenet competition, Microsoft came up with a deep learning model of 8 layers as a solution for stegnalysis. Ever since then, deep learning has been the preferred method for carrying out stegnalysis.

II. RELATED WORK

In [1]The authors propose a mechanism to use unpredictable distortion matrices for data hiding. The cover images are then grouped into subsequences, each coded with a static payload. Finally, a spatial-rich-model stegnalyser is used with an ensemble classifier and the process is carried out over a very large sample set to minimize loss. In [2] however, image foresnics are used to separate processed and investigated images before stegnalysis exploiting the fragility of image manipulation detection to reduce the false alarm rate by fragile detection of gamma transformation and LSB matching. In [3], the authors propose the removal of MSB planes as statistics suggest that embedding is usually not carried out in these planes. Further, since the image actually shrinks after data embedding, the authors propose the use of diagonal elements of the co-occurrence matrix to reduce the effect of asymmetry caused in the same due to data embedding. In [4]image net model is used to carry out stagnalysis transfer learning constructed using a double pipeline architecture – for cross-entropy and the optimization through a 10-weight decay. The hyper-parameters for each of the sample was calculated and the model was executed on the ALSAKA-II dataset. In [5], a universal stenalysis engine is proposed - a system that holds all the known steganlysis models and then based on the cover image analysis, a suitable algorithm,a suitable classification model for feature extraction and thus a new feature set classifier is modelled and trained. The authors in [6] propose preprocessing to be carried out in CNN based image processing because under reduction, the frequency for certain values becomes too high. This involves the convolution of reduced stegno images only through a trained model instead of convoluting both reduced stegno images and reduced cover images. This helps in selective increase in pixel value frequencies. In [7],proposal is for the utilization of deep neural networks to train a model that can distinguish between stegno images and normal images, gener-ating two symmetrical subnets similar to [4] in three phases. This creates an adaptive stegnalysis model that extracts a hyper parameter through training and classification. In [8], through the use of Deep learning with ensemble learning, the authora proposes that a contournet transformation be used based on a statistical analysis of the pixels. The ROC model is utilized and accuracy is computed using a training model. The authors finally conclude that while on grayscale images, such a model can yield high accuracy, it is not exactly the case with colourful images. [9] presents a novel framework with three parts. The first has four parts – shallow feature extraction, un-looped feature extraction, deep feature reduction and classification. The second and third involve residual learning models. The framework reduces memory dependency and processor usage. The results yield a new baseline architecture that can be used to detect a new hyper-parameter. However detection for a new hyper-parameter increases the computational cost of the module. The authors conclude that stegnalysis models do not very accurately determine encrypted images and rather use heuristics to detect noise in images. In [10], suggestion is made over several methodologies and tricks to carry out both pre-processing and stegnalysis, including the creation of a model that directly inputs the DCT coefficients, further aided by a multi-layer model, pseudo labelling of features, stacking those features with a tree – all this built upon a CNN based model. [11] proposes a model which involves an ensemble of several filters and tools – a binary relevance multi-label filter followed by a Logistical Regression classifier and finally a margin FML - to ensure that the model uses the right set of

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 08, August-2021

images for classification and stegnalysis, reducing false positives. In [12], the authors propose steganalysis based on LSB matching by using Ye network (YeNet). The YeNet produces a 2-D vector output which are the dimensions of the features extracted. A linear model is compared to the dimensions thus obtained and the origin is treated as a critical point. This is used to classify the image as a stego or a cover image. The experiments suggest that in YeNet, the output of stego image changes regularly with the embedding rate, the mapping relationship between embedding rate and feature distribution can be fitted, so as to estimate the embedding rate of a given stego. [13] states that during stegnalysis, the system should use diagonal elements of the matrix just as suggested in [3] to eliminate noise. It also suggests the use of Gaussian elimination equations. Then a probability model is used to distinguish between cover images and stegno images. The authors conclude that while there is marginal noise in the restored image as well, however, in smaller sections of the image, if there is no correlation between adjacent pixels, a probability function is again employed to detect an absorber edge and embedded data. [14] points out that noises can be eliminated by splitting exposure time and then averaging the images under several captures making it useful for images under motion, which are seemingly blur. This is restored by splitting the motion under several captures. Retrieve a normalized correlation coefficient. Finally, an average from all the normalized pixels is obtained. This reduces blurring as well as noise. [15] is similar to [14] where, a novel methodology that is useful for reduction in time complexity and improvement in error rate is proposed. This is achieved through introducing Legrange multipliers over a blurred image modelled like an object in linear motion. This transforms the image into an equation of matrices and finally a restoration algorithm is applied. This restoration algorithm is also based on legrange multiplier. [16] makes a comparison between inverse filtering and Wiener filtering on blurred images. The authors conclude that both the filters work perfectly under conditions where images are blurred and have minimal, marginal or no noise. However, in the presence of noise, a new methodology is needed. In [17], authors propose several steps of filtering – beginning with median filtering, wiener filtering, regularization filtering, finally followed by Lucy-Richardson filtering. This is then measured using several metrics like the PSNR, MSE and SSIM values..

III. DATASET

1. The ALASKA-II dataset because it is has a set of unique elements, has yielded very poor rates of false positives. It is also a fresh dataset released in 2020.
2. The naıve image dataset present to retest the results obtained on the first dataset.

IV. PROPOSED WORK

The model shall be facilitated with a LSB filter as proposed in [3] and [12].

A. Aim of Experiment

The aim behind different models were different. As discussed, in the model with Transfer Learning, the dataset should quickly adapt to the base stenographic functions used. While, without Transfer Learning, the purpose was to make the algorithm perfect to identify a family of steganography algorithms and not to all steganography functions.

B. Software and Harware Requirements

1) Python 3.9.2
2) TPU Accelarator
3) ALASKA2 Dataset

The author used the Kaggle Jupyter Notebook Platform to achieve these.

C. Supposed Challenges

Increasing network depth does not work by simply stacking layers together. Deep networks are hard to train because of the vanishing gradient problem — as the gradient is back-propagated to earlier layers, repeated multiplication may make the gradient extremely small. As a result, as the network goes deeper, its performance gets saturated or even starts degrading rapidly.

V. EXPERIMENTAL RESULTS

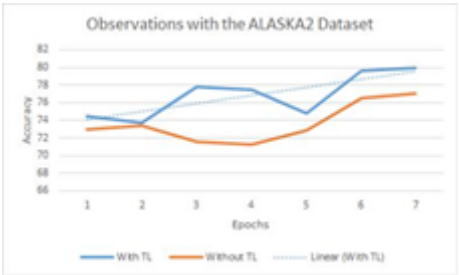

figure: Observations with ALASKA-II dataset (blue lines show Transfer learning accuracy vs epochs, orange line shows non-TL model)

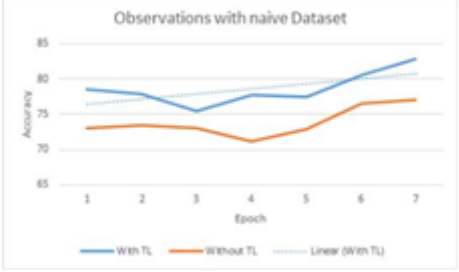

figure: Observations with the naïve dataset (blue lines show Transfer learning accuracy vs epochs, orange line shows non-TL model)

VI. OBSERVATION

The following results can be arrived upon based on th experimental observations recorded with both the data-sets:

1) As already demonstrated in Experimental Observations, the accuracy and loss in both Transfer Learning based model was 77percent accuracy compared to 58 percent loss.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 08, August-2021

2) The transfer learning based model that uses two epochs has a loss of 63 percent as compared to the 58 percent loss in the same model but with a single epoch of training. This might be because Deep Learning models approach a saturation when the same model undergoes various levels of training. This was one of the reasons why I opted for RESNET-50 over RESNET-150 or higher. The Transfer Learning paradigm only increases the effective number of layers in the Deep Learning Model and is slower because a more comprehensive memory of the layers has to be maintained by the algorithm.

3) Without Transfer Learning, the accuracy was around 74 percent for RESNET – 50 model with ALASKA2 dataset and around 75 percent for Na¨ıve Dataset.

4) In the Naive dataset that I used once the model was successfully trained with ALASKA2 dataset, the accuracy achieved was about 78 percent compared to a loss statistics of around 58 percent during model training. This is when the number of epochs is 1.

5) The Naive Dataset training was faster and slightly better results were obtained as compared to ALASKA2 dataset, probably because the Naive Dataset is DCT encoded and LSB is a simple and yet effective way to decode the same.

6) The conclusions are based on the two datasets that I used to train and test the model. This might however be specific to the datasets I picked up because the accuracy largely depends upon the number of steganography algorithms used in the data studied under the model as well. This is because the Transfer learning model will remember basic algorithms when trained while the non-transfer learning model will learn some specific families of algorithms better.

VII. CONCLUSIONS

The author concludes that the algorithm which includes a simple LSB Filter put before a RESNET-50 model trained with Transfer Learning Paradigm has an average accuracy score of 75 percent with the average training losses amounting to 60 percent. And the transfer learning model clearly outweighs the model without transfer learning as in most cases, the user will not run over an infinitely large number of epochs in any deep learning model.

VIII. SUGGESTIONS AND FURTHER IMPROVEMENTS

Using the Xception paradigm(Extreme Inception) model that was introduced yet again for the ImageNet 2015 dataset might help to further improve the RESNET+LSB model accu-racy. The Xception model outperformed the then best model (before the ImageNet 2015 competition where Microsoft pro-posed the RESNET model) by eliminating the need to learn 3D mapping and replacing it with learning a 2D+1D map. The model might also be improved further with accuracy of about 85 percent achieved if the CNN model being used in the RESNET-50 is replaced with a DenseNet Model. Densenet is a Recurrent CNN model with added memory. However, I did not suggest it for two reasons – it will make the model very complicated.

1. The Densenet model itself behaves as a small deep learning paradigm. In this case, several Densenets might have the same effect on the model as increasing the number of deep learning layers or transfer learning layer has – increased saturation in training and increase loss.
2. Additionally, the RESNET model is experimentally known to be the strongest Stegnalysis mode. However, the researcher is open to the idea of combining these two models for the sake of experimental verification of point (2).

ACKNOWLEDGMENT

The author expresses his gratitude to Prof Sureshkumar N, Associate Professor, School of Computer Science, VIT, Vellore for being a teacher and guide throughout this project and showing his confidence in the feasibility of the same.

REFERENCES

[1] Jyoti Neginal , and Dr Ruskar Fatima, ”An adaptive stenographic technique, ascertaining upper bounds of embedding” IEEE,2019
[2] Ping Wang, and others, ”Steganalysis aided by fragile detection of image manipulations” IEEE,2019
[3] M. Abolghasemi, and others, ”Steganalysis of LSB Matching Based on Co-Occurrence Matrix and Removing Most Significant Bit Planes”IEEE, 2008
[4] Yassie Yousfie, and others , ”ImageNet Pre-trained CNNs for JPEG Steganalysis” IEEE, 2020
[5] Yanping Tan, and others, ” New Design Method about the Universal Steganalysis Classifier in Digital Image” IEEE, 2020
[6] Hiroya Kato, and others, ”A Preprocessing Methodology by Using Additional Steganography on CNN-based Steganalysis” IEEE, 2020
[7] Wieke You, and others, ” Siamese CNN for Image Steganalysis” IEEE, 2021
[8] Nour Mohamed, and others, ”A Review of Color Image Stegnalysis in the Tranform Domain” IEEE, 2020
[9] Wonhyuk Ahn, and others, ”Local-Source Enhanced Residual Network for Steganalysis of Digital Images” IEEE, 2020
[10] Kaizaburo Chubachi, ”An Ensemble Model using CNNs on Different Domains for ALASKA2 Image Steganalysis” IEEE, 2020
[11] Tayebe Abazar, and others, ”A binary relevance adaptive model-selection for ensemble stegnalysis” 17th International ISC Conference on Informa-tion Security and Cryptology (ISCISC) 2020
[12] Yu Sun, and Taiyun Li, ”A Method for Quantitative Steganalysis Based on Deep Learning” IEEE, 2019
[13] Varlen Grabski, ”Digital Image Restoration Based on Pixel Simultaneous Detection Probabilities” IEEE, 2009
[14] Zile Wei, ”Digital Image Restoration by Exposure-Splitting and Registration” IEEE 2004
[15] Igor Stojanovic, ”Application of Non-Iterative Method in Digital Image Restoration” IEEE
[16] Mohammed Mahmudur Rahman Khan, and others, ”Digital Image Restoration in Matlab: A Case Study on Inverse and Wiener Filtering” IEEE 2018
[17] Reecturaj Mishra, and others, ”Digital Image Restoration using Image Filtering Techniques” IEEE 2019
[18] Kiaming Hu, and others, ”Deep Residual Learning for Image Recognition” arXiv:1512.03385v1 [cs.CV] 10 Dec 2015