2018 11th International Conference on Developments in eSystems Engineering (DeSE)

Steganalysis of RGB Images Using Merged Statistical Features of Color Channels

Zaid I. Rasool
Dept. of Computer Technology Eng.
Al-Mustaqbal University College
Babel
Iraq
Zaid.ibrahim@mustaqbal-college.edu.iq

Mudhafar M. Al-Jarrah *
Dept. of Computer Science
Middle East University
Amman
Jordan
maljarrah@meu.edu.jo

Saad Amin
Dept. of Information Technology
Al-Khawarizmi Int. College
Abu Dhabi
UAE
saad.amin@kic.ac.ae

***Abstract*—This paper presents a steganalysis model that uses an enhanced grayscale statistical feature set, in the detection of data hiding in uncompressed RGB color images. A dataset of 3000 RGB images is created, using natural images from public sources, in TIFF and JPEG formats, that are converted to BMP format and resized to 512x512 pixels. The clean images are embedded with secret image data, using two payload schemes, 2 bits per channel (bpc) and 4 bits per channel. The selected feature set consists of 24 features per color channel, 72 features per image, which includes the Gray Level Co-Occurrence Matrix (GLCM) features, Entropy features, and statistical measures of variation. The feature set elements are calculated for individual channels, combined into image features vector. The steganalysis process is based on supervised machine learning, utilizing the Support Vector Machine (SVM) binary classifier's implementation in MATLAB. The results show very high detection accuracy for the two cases of 2-bpc and 4-bpc embedding schemes. Also, there are no noticeable differences in the detection accuracy between the two sources of images, even though un-compression of the JPEG images has reduced their noise contents. The paper ends with a conclusion and suggestions for future work.**

Keywords—steganalysis, feature set, GLCM, entropy, machine learning, SVM classifier, detection accuracy, RGB, bit per channel

I. INTRODUCTION

The work in this paper deals with the problem of detecting the presence of hidden data inside uncompressed color images. In particular, the work presents an enhanced scheme for the detection of hidden secret data within RGB (Red, Green, and Blue) cover images, based on statistical textural features, and by using machine learning steganalysis techniques.

II. RELATED WORK

Steganalysis of 8-bit grayscale images has been the subject of extensive research [1-5], which has resulted in models and feature sets that are based on analyzing textural features of images. In [6], a feature set is presented which combines GLCM features [7], Entropy, and other statistical measures of variation. The features are directed to measure properties of the right-half of images' bytes. The steganalysis of color images has been applied to compressed images such as JPEG [8], and uncompressed images [9]. Research on feature-based steganalysis of color images has resulted in feature sets that are based on attributes of the color channels, and the relationships between channels, such as the Color Gradient Co-Occurrence Matrix (CGCM) model [10, 11], and the Color Rich Model [12].

III. PROPOSED MODEL

The proposed steganalysis model aims to detect the presence of a hidden message within an uncompressed RGB cover image. The detection task is based on prior training of a binary classifier on features of a dataset of clean and stego images, using supervised learning techniques. The steganalysis approach is based on merging features of single color channels into a multi-channel feature set, without consideration to the correlation between color channels. The detection accuracy of the proposed model is evaluated using a dataset of uncompressed RGB clean images, and the associated stego images, with embedding rates of 2 bits per channel and 4 bits per channel.

IV. FEATURES SELECTION

The statistical textural features of the proposed model consist of two parts: a single channel feature set and a multi-channel feature set. The single channel feature set includes standard texture features as well as additional statistical features. The proposed model is

978-1-5386-6712-5/18/$31.00 ©2018 IEEE
DOI 10.1109/DeSE.2018.00048

founded on a channel-based feature set that is merged into a 3-channel feature set for image-based steganalysis. The channel-based feature set (CFS) consists of GLCM features (Contrast, Correlation, Energy and Homogeneity), as well as other textural features such as Entropy, which was suggested by Haralick [7] in the study of textural features of images, and have been used in many steganalysis research works [6]. The channel-based feature set consists of the features shown in Table I.

TABLE I. SINGLE CHANNEL FEATURES

Feature Name	Feature Description
CC-LR	Correlation coefficient between left and right half-bytes
CoV-FB	Coefficient of variation of full-bytes
CoV-RHB	Coefficient of variation of right half-bytes
GLCM-FB	Contrast, Correlation, Homogeneity, Energy, of full-bytes
GLCM-RHB	Contrast, Correlation, Homogeneity, Energy, of right half-bytes
GLCM-3LSB	Contrast, Correlation, Homogeneity, Energy, of 3LSB part of byte
GLCM-2LSB	Contrast, Correlation, Homogeneity, Energy, of 2LSB part of byte
Entropy-FB	Entropy of full-bytes
Entropy-RHB	Entropy of right half-bytes
Skew-FB	Skewness of full-bytes
Skew-RHB	Skewness of right half-bytes
AveDiff-RHB	Average difference between successive right half-bytes

V. THE DATASET

The selected cover image type is uncompressed RGB-BMP, in three channels, without the alpha channel. Two independent datasets are used, for double validation of the proposed model. The first validation dataset consists of 1500 clean images in TIFF format with alpha channel, that were downloaded from the Natural Resources Conservation (NRC) image dataset [13]. The original NRC images were converted from TIFF to BMP format and resized to 512×512 pixels. The second validation dataset is based on the CALTECH's birds images dataset [14], which is in a compressed color JPEG format. A set of 1500 CALTECH images were converted to BMP format and resized to 512×512 pixels.

Fig. 1 and Fig. 2 show a sample of the converted and resized NRC and CALTECH images. The two datasets are available for download in [15].

Fig. 1. Sample NRC cover image

Fig. 2. Sample CALTECH cover image

The clean images from the two datasets were embedded with secret data of different sizes, to generate the stego datasets. Three images were used as the secret data; a large image for high capacity embedding using the 4-bpc scheme, a medium image for low capacity embedding using the 2-bpc scheme, and a small image for single channel embedding using the 2-bpc scheme.

The secret data for the 3-channel high payload embedding with 4-bpc is the image "House.bmp" from USC-SIPI Image Database [16], as shown in Fig. 3. It was resized to fit the maximum embedding capacity of the selected cover images at a payload of 50% of the cover image's size.

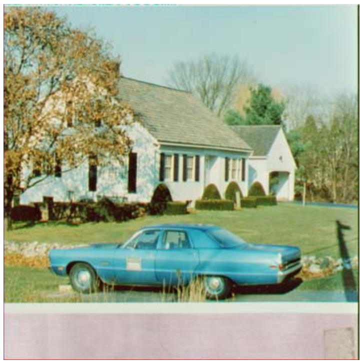

Fig. 3. Large secret image House.bmp, 360×360, 379 KB, 50% payload

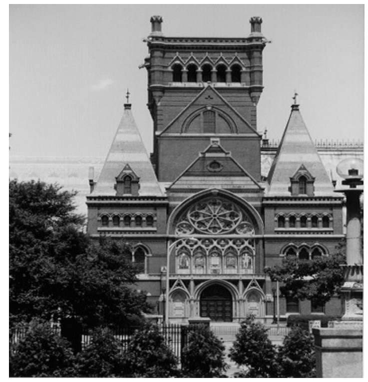

Fig. 5. Small secret image Harvard.jpg, 354×520, 63 KB, 12.5% payload

For the 3-channel low payload embedding with 2-bpc, the image "Peppers.bmp" from the Gonzales dataset [17] is used, as shown in Fig. 4, which was resized to fit the maximum embedding capacity of the selected cover images at a payload of 25% of the cover image's size.

Fig. 4. Medium secret image Peppers.bmp, 254×254, 189 KB, 25% payload

For single channel embedding with 2-bpc, the image "Harvard.jpg" (Wikimedia.org images) is used, as shown in Fig. 5. The original image was resized to 63 KB, to fit the maximum embedding capacity of the selected cover images in a single channel, at a payload of 12.5% of the cover image's size.

VI. EXPERIMENTAL WORK

The research methodology is based on an experimental evaluation of the proposed model that involved embedding, features vectors extraction from the clean and stego images, and machine learning classification of the extracted features vectors, using 3-fold cross-validation [18].

A. Embedding

The clean images of the two datasets were embedded with secret data using the spatial domain steganography. For the 3-channel embedding, the pixels of each image were embedded with secret data sequentially, where each channel in each pixel were embedded with 2 bits or 4 bits by replacing the least significant bits. For single channel embedding, only the NRC cover images were used, in which the Blue color channel of each pixel was embedded using 2-bpc. The processes of embedding have produced five stego datasets: NRC-LSB2, NRC-LSB4, CALTECH-LSB2, CALTECH-LSB4, and NRC-2LSB-Blue.

B. Features Extraction

The selected channel-based features are extracted from each color channel of the clean and stego images. The features are extracted using built-in functions in MATLAB, including GLCM, Entropy, Correlation Coefficient, Standard deviation, Mean, and Skewness. The coefficient of variation (CoV) is calculated as below:

$$CoV = \frac{\text{Standard Deviation (n)}}{\text{Mean (n)}} \quad (1)$$

where n is a vector of single-channel full-bytes or parts of bytes of a single channel.

The output from the feature extraction process is a features data file in Excel CSV format that contains features of single channels in separate worksheets. The single channel worksheets are combined into an image features data file which consists of 72 feature elements. The image features file are divided into training subset and testing subset, according to the 3-fold cross-validation method.

The training file is formed by merging equal number of clean and stego images' features data. Similarly, the testing (unseen) files contain features data of images that are not part of the training data. For each cross-validation cycle, two-thirds of the features data vectors are used for labeled training, and the other third is used for testing.

C. Classification

In this phase, the Support Vector Machine (SVM) classifier is chosen for the binary classification task of images, due to its superior performance in pattern and image processing applications [19]. The MATLAB implementation of the binary SVM algorithm is utilized, using the "svmtrain" and "svmclassify" functions. The training subset of features data vectors of a batch of labeled 1000 clean and 1000 stego images were used for training the "svmtrain" function. Based on the training outcome, the unlabeled testing subset of the features data vectors of 500 clean and 500 stego images were analyzed using the "svmclassify" function, where each unlabeled testing image is classified as either "clean" or "stego". The training / testing process was repeated three times using a 3-fold cross-validation approach. The two datasets (NRC and CALTECH) were processed separately, accordingly the detection results were independently computed.

D. Evaluation Metrics

The following metrics have been used in evaluating the detection performance of the proposed model:

- True Negative (TN): The ratio of true negative detections to the number of clean images.
- True Positive (TP): The ratio of true positive detections to the number of stego images.
- False Negative (FN): The ratio of false negative detection to the number of stego images.
- False Positive (FP): The ratio of false positive detection to the number of clean images.
- Detection Accuracy: The ratio of correctly detected clean and stego images to the total number of clean and stego images, as below:

$$\text{Detection Accuracy} = \frac{(TN+TP)}{(TN+TP+FN+FP)} \qquad (2)$$

VII. RESULTS AND DISCUSSION

The batch classification of the clean and stego images of the two datasets, using 3-fold cross-validation, resulted in error rates of FP, PN, TP and TN, as well as the overall detection accuracy. Table II shows the error rates and detection accuracy results of the steganlaysis of the NRC dataset, using 2-bpc and 4-bpc embedding. The results show very high detection accuracy for both the 2-bpc and 4-bpc cases, despite the big difference in the size of the embedded data. Also, the false negative (FN) and false positive (FP) results are identical, which indicates that the proposed model has a balanced detection of both errors.

TABLE II. RESULTS OF THE 3-FOLD CROSS-VALIDATION OF RGB FEATURES OF THE NRC DATASET

Metric	**Average of 3 folds (%)**	
	2-bpc	***4-bpc***
FN	0.07%	0.00%
FP	0.07%	0.00%
TN	99.93%	100.00%
TP	99.93%	100.00%
Accuracy	99.93%	100.00%

Table III shows the error rates and detection accuracy results of the steganlaysis of the CALTECH dataset, using 2-bpc and 4-bpc embedding. The results show a similar pattern to the NRC results, despite the fact the CALTECH dataset images were converted from lossy compressed images, thereby having less noise than the NRC images that were converted from lossless TIFF images.

TABLE III. RESULTS OF THE 3-FOLD CROSS-VALIDATION OF RGB FEATURES OF THE CALTECH DATASET

Metric	Average of 3 folds (%)	
	2-bpc	***4-bpc***
FN	0.47%	0.00%
FP	1.33%	0.07%
TN	98.67%	99.93%
TP	99.53%	100.00%
Accuracy	99.10%	99.97%

Table IV shows the error rates and detection accuracy results of the steganlaysis of the NRC dataset, using 2-bpc embedding in the Blue channel, without alteration to the Red and Green channels. The results show a similar pattern to the 3-channel embedding, which indicates that the proposed model and the classifier performed equally well even when only one channel features were altered.

TABLE IV. RESULTS OF THE 3-FOLD CROSS-VALIDATION OF RGB FEATURES OF THE NRC DATASET WITH BLUE-CHANNEL EMBEDDING

Metric	Average of 3 folds (%)
	2-bpc in the Blue channel
FN	0.07%
FP	0.53%
TN	99.47%
TP	99.93%
Accuracy	99.97%

VIII. CONCLUSION

The work in this paper presented a steganalysis model to detect the presence of hidden data inside uncompressed RGB color images, using statistical textural features that are based on properties of grayscale 8-bit images, which include GLCM, Entropy and statistical measures of variation. The selected features were extracted from datasets of clean and stego images, and classified using the Support Vector Machine algorithm. The focus of this work is in analyzing the individual color channels features separately, then to merge features of the three channels into a global image feature set.

Cross-validation of the steganalysis results of two independent datasets of clean and stego images showed very high detection accuracy, over 99% for both cases of 4-bpc and 2-bpc embedding schemes. Also, single channel embedding yielded similar high detection accuracy.

For future work, the proposed steganalysis model can be evaluated using lower embedding rates and different media types such as audio and video. The proposed model can be extended to include additional features to deal with other steganography models such as the transform domain.

REFERENCES

[1] A. Cheddad, J. Condell, K. Curran and P. McKevitt, "Digital image steganography: survey and analysis of current methods," Signal Processing, 90(3), 727-752, 2010.

[2] K. L. Chiew, "Steganalysis of binary images," PhD thesis, Macquarie University, 2011.

[3] A. Aljarf, S. Amin, J. Filippas, and J. Shuttelworth, "Develop a detection system for grey and colour stego images," International Journal of Modeling and Optimization, 3(5), 458, 2013.

[4] A. Ker, "Steganalysis of LSB matching in grayscale images," IEEE Signal Processing Letters 12(6):441-444, 2005.

[5] D. Lerch-Hostalot and D. Megıas, "LSB matching steganalysis based on patterns of pixel differences and random embedding," Computers & Security, Volume 32, Pages 192- 206. February 2013

[6] M. M. Al-Jarrah, Z. Al-Taei, and A. Abuarqoub, "Steganalysis Using LSB-Focused Statistical Features", Proceeding, ICFNDS '17, proceedings of the International Conference on Future Networks and Distributed Systems, 2017.

[7] R. M. Haralick, K. Shanmugam, and I. Dinstein, "*Textural features for image classification*," IEEE Trans. on Systems, Man and Cybernetics, 1973.

[8] V. Holub, and J. Fredrich, "Low-complexity Features for JPEG steganalysis using undecimated DCT," IEEE Transactions on Information Forensics and Security (Volume: 10, Issue: 2, Feb. 2015.

[9] H. Abdulrahman, M. Chaumont, P. Montesinos, B. Magnier, "Color images steganalysis using RGB channel geometric transformation measures," Security and Communication Networks, 2015.

[10] R. Gong, and H. Wang, "Steganalysis for GIF images based on colors-gradient co-occurrence matrix." Optics Communications, (vol. 285), no. 24, pp. 4961-4965, 2012.

[11] A. Aljarf, S. Amin, J. Fillipas, and J. Shuttelworth, "The development of an image detection system based on the extraction of color gradient co-occurrence matrix (CGCM) Features," Dese2016 Conference, UK, 2016.

[12] M. Goljan, J. Fridrich, and R. Cogranne, "Rich model for steganalysis of color images," IEEE International Workshop on Information Forensics and Security (WIFS) (pp. 185-190). IEEE, 2014.

[13] NRCS Photo Gallery, natural resource, and conservation related photos, https://photogallery.sc.egov.usda.gov/res/sites/photo gallery, viewed on 1/2/2017.

[14] CALTECH-UCSD Birds-200-2011 Dataset, http://www.vision.caltech.edu/visipedia/CUB-200-2011.html, viewed on 1/2/2017

[15] M. Al-Jarrah, "RGB-BMP Steganalysis Dataset", Mendeley Data, v1, 2018 http://dx.doi.org/10.17632/sp4g8h7v8k.1

[16] USC-SIPI Image Database, http://sipi.usc.edu/database/, viewed on 1/3/2017

[17] R. C. Gonzales, and R. E. Woods, Digital Image Processing, 3rd Ed.,Pearson, image databases, http://www.imageprocessingplace.com/root_files_V3/image_databases.htm, viewed on 1/3/2017

[18] Y. Abu-Mostafa, and M. Magdon-Ismail, Learning from data, AMLBook, 2012.

[19] A. Tzotsos A. and D. Argialas, "Support Vector Machine classification for object-based image analysis," Object-Based Image Analysis. Lecture Notes in Geoinformation and Cartography, Springer, Berlin, Heidelberg, 2008.