International Journal of Computer Trends and Technology
ISSN: 2231 – 2803 / https://doi.org/10.14445/22312803/IJCTT-V70I5P102

Volume 70 Issue 5, 7-14, May 2022

Original Article

Security Test Using StegoExpose on Hybrid Deep Learning Model for Reversible Image Steganography

Awodele Oludele[1], Idowu Sunday[2], Kuyoro Afolashade[3], Nzenwata Uchenna[4]

[1,3,4]*Computer Science Department, Babcock University, Ilishan-Remo, Ogun State, Nigeria.*
[2]*Software Engineering Department, Babcock University, Ilishan-Remo, Ogun State, Nigeria.*

Received: 11 March 2022 Revised: 30 April 2022 Accepted: 04 May 2022 Published: 21 May 2022

Abstract *- Image steganography is an act of concealing secret information using the image as a cover medium. It is said to be reversible when the same level of importance placed on the retrieval of the secret information is also placed on the recovery of the cover image. The process of hiding information in a cover is called steganography while retrieving the information that was hidden using steganography is called steganalysis. Image steganography is faced with challenges in payload capacity, security, and robustness. Attempts have been made to bring a good solution to this problem but end with a trade-oof in the payload capacity and the security. This paper attempts to solve this problem by proposing a Hybrid Deep Learning Model, which comprises DNN, CycleGAN, and CNN deep learning tools. The study's outcome yielded a good payload capacity and a good security measure, which was evaluated using PSNR and SSIM.*

Keywords *- SSIM, CycleGAN, Payload, Security, StegoExpose.*

1. Introduction

According to [1], image steganography is an act of concealing secret information using the image as a cover medium. It is said to be reversible when the same level of importance placed on the retrieval of the secret information is also placed on the recovery of the cover image. The process of hiding information in a cover is called steganography while retrieving the information that was hidden using steganography is called steganalysis [2]. Image steganography is faced with challenges in payload capacity, security, and robustness [3]-[5]. Attempts have been made to bring a good solution to this problem but end with a trade-off in the payload capacity and the security [6], [7].

In addressing this problem, previous studies have used varieties of image steganography techniques such as Spatial Domain Transform (SDT) [8], Frequency Domain Transform (FDT) [9], Convolutional Neural Network Steganography (SteganoCNN) [10], and Generative Adversarial Network Steganography (SteganoGAN) [11]. The SDT and SteganoGAN yielded high payload capacity image steganography systems. However, they ended up with deformed stego-images, which is an indication that the security of the system is poor [12]. Contrary to these approaches, the Frequency Domain and SteganoCNN focused solely on addressing the system's security without considering the payload capacity. Therefore, there is a trade-off between payload capacity and security among the existing models. For this reason, there are no existing image steganography systems that assure good payload capacity and security, which remains a problem that should be addressed.

This paper is an extract from a major work where the development of reversible Image steganography using the Hybrid Deep Learning Model. The model comprises a cover selection model using the basic Deep Neural Network, the encoding model using Cycle-consistent Generative Adversarial Network (CycleGAN), and an adversarial Convolutional Neural Network for the decoding model. This paperwork aims to show some extracted results obtained from the complete study and show how the security of the study was ascertained using stegoExpose steganalysis tools.

2. Features of Reversible Image Steganography

2.1. Imperceptibility

The greatest priority criterion for any data embedding is imperceptibility, as the fundamental characteristic and strength of any steganographic approach are hiding the hidden data in the digital image so that it cannot be grasped by the naked human eye or statistical methods.

2.2. Security

In a steganographic system, security refers to unnoticeability or undetectability. As a result, any steganography technique is considered secure if the secret

data is not detectable by statistical means or is removed after being detected by the attacker. The secure transmission of secret data is a key requirement of the steganographic process. As a result, security is the primary concern to prevent unauthorized persons or computers from accessing data transmitted over an open channel.

2.3. Payload Capacity

An effective steganographic system always attempts to convey as much information as possible while utilizing as little cover material as possible. This reduces the possibility of interception while communicating across an insecure network and, as a result, generally necessitates a large embedding capacity.

2.4. Robustness

This indicates the embedding and decoding scheme's capacity to function even if the stego-image is damaged by a third-party using image processing techniques such as rotation, scaling, resizing, etc.

3. Steganalysis

Steganalysis is the study that understands the analysis of steganography. It employs the techniques that are used to retrieve secret communication from stego-objects. In [13], steganalysis is likened to the hypothesis-testing problem because a steganalyst will want to know whether or not a certain cover is a stego-object. This is not far-fetched from the idea shared in [14], where steganalysis was treated as a classification problem using machine learning algorithms. In recent times, deep learning techniques have been used for similar steganalysis operations. We do not intend to address existing classes of steganalysis in this study. Therefore, the intent is to discuss applicable steganalysis tools in this study. The most commonly used steganalysis algorithms are;

The Ensemble Classifiers are used to enhance the accuracy of predictive analytics and data mining applications. The ensemble classifiers operation runs concurrently, closely related, but with different modeling analytics, and at the end of the execution, the outcomes are brought together as a single output. This steganalysis method was used in [15] for the steganalysis of digital media.

The Regularised Linear Classifier utilizes Support Vector Machine (SVM) classifiers, a binary classification machine learning technique [16]. This classifier uses its sub-classifiers, which are learned by the fusing SVM classifier, to perform steganalysis. In this case, features from the cover and stego-object are extracted, and the extracted features are sorted into groups based on feature correlation. The detection findings are used to train the fusion classifier.

Convolutional Neural Networks (CNN-based steganalysis) is the state-of-the-art steganalysis method to reveal secret communication using a deep neural network. Due to the powerful classification algorithm, CNN-based steganalysis can also reveal GAN-based steganography using a variant of itself, a deep convolutional neural network (DCNN).

StegoExpose is a steganography detection program that detects steganography in images. It includes a command-line interface intended to analyze images in batch while also giving reporting and customization tools understandable to non-forensic specialists. The StegExpose grading system is based on an intelligent and fully proven mix of pre-existing pixel-based steganalysis algorithms. Apart from identifying steganography in images, StegoExpose also helps determine the hidden message's length.

In image steganography, ensemble and linear classifiers are both image steganalysis approaches. As a result, one of the methods used to assess the security of steganography models is the stegoExpose [17], a conventional steganalysis tool based on the ideas of ensemble and linear classifiers.

4. Methodology

The conceptual idea behind reversible image steganography is given in figure 1, where cover images are selected appropriately based on the secret images that we want to hide. This is handled by the encoder or hiding model. the output of this model produces the stego-image, which is fed into the decoder model or the extraction model. The outputs are the reconstructed cover image and the reconstructed secret image.

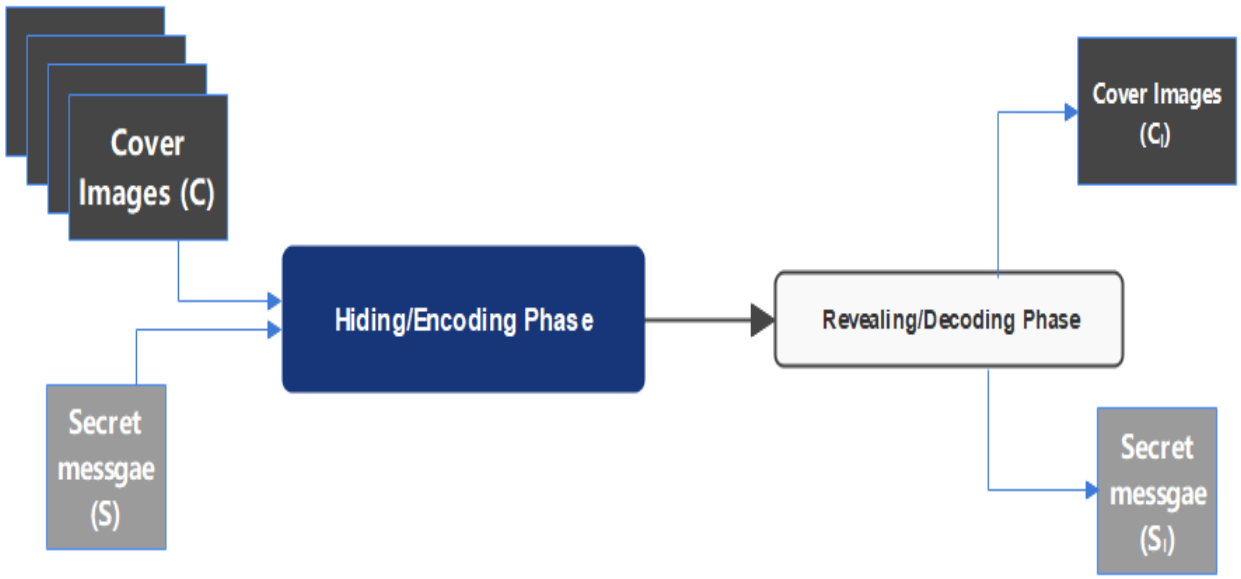

Fig. 1 Reversible image steganography concept

The methodology adopted in the general study is similar to what is obtainable using the concept of image steganography. The difference is that Fig 1 did not consider the data preparation phase and cover selection model. Figure 2 summarizes the general proposed model shown. There are three basic phases involved in the study.

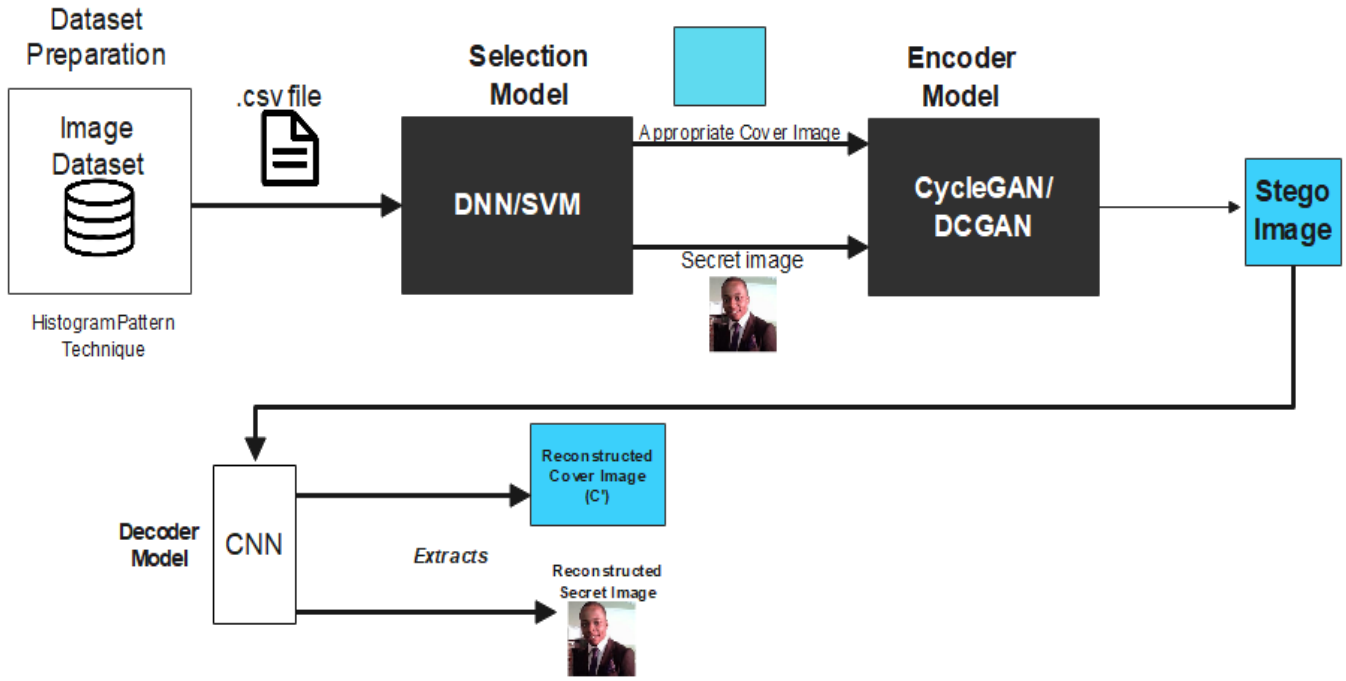

Fig. 2 The proposed hdlm.

The first phase has to do with preparing the data and developing the cover selection model. The encoding phase model is the second phase, where the secret image is embedded using the cycle gang. The third phase is the decoding model using CNN.

4.1. Data Preparation

The Microsoft Common Objects in Context (COCO) unlabeled dataset of over 200 000 images, dated 2017 [18], was used for this study. The data was prepared to ease model training. To ease the model training using the unlabeled image dataset, the Histogram of Pattern Sets (HoPS) technique as used in [19] was adopted. The outcome of this technique generated a comma-separated value (csv) file, which served as a look-up file to the image repository. It was used to train the algorithm that classifies suitable cover images and the corresponding secret image.

4.2. Cover Selection Model

The cover selection is a classification problem that uses the HoPS-generated details as parameters. The cover image selection is part of the preparation step. This was accomplished by training a selection model, which aided the encoding preparation. This step aims to assist in selecting the best cover image that will be most appropriate to conceal the secret image, as selecting an appropriate cover image is essential in deciding the efficacy of a steganographic system. This study used the deep learning architecture known as the Deep Neural Network (DNN) [25] for cover image selection.

4.3. Encoder Model

The encoding phase is the concealing network that creates the stego-images or carrier images. This study proposes the Cycle-Consistent Generative Adversarial Network (CycleGAN) [25] to achieve this phase. The encoding model receives cover images and the corresponding secret images from the cover selection model. The cover image selection model (DNN model) is used alongside the secret image at the CycleGAN model's generator G for the encoding process. After a successful encoding, the stego image is discriminated with the cover image using the CycleGAN model's discriminator (D). This continues until the discriminator fails to discriminate between the original cover image and the stego image. At this point, the stego image is sent as output from the encoding phase. The general GAN's loss equation described in equations 1 and 2 by [20], [21] was used to optimize the losses in the encoder's network.

$$\text{Loss} = \text{Min}_{(G)}\text{Max}_{(D)}[\log(D_{(x)}) + \log(1\text{-}D(G_{(z)}))] \quad \ldots\ldots\ldots \quad 1$$

Equation 1 was used by considering a single data point. To consider the entire data set, equation 1 is transformed to equation 2

$$\text{Min}_{(G)}\text{Max}_{(D)}\, V(D, G) = \text{Min}_{(G)}\text{Max}_{(D)}(E_{x\sim Pdata(x)}[\log(D_{(x)})] + E_{z\sim Pdata(z)}[\log(1\text{-}D(G_{(z)}))]) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 2$$

The stego image was analysed to obtain the payload capacity and the security of the steganography system.

4.4. Decoder Model

The decoding phase is also referred to as the revealing network, where the stego-image is decoded using an adversarial Convolutional Neural Network (CNN). After that, the secret message is revealed, and the reconstructed cover image is obtained. Beyond extracting the secret message, the reconstructed cover image from the CNN decoder was compared with the original cover image. The steganography system is considered to pass the reversibility test if the outcome passes the Human Vision System by measuring the Peak Signal-to-Noise Ratio (PSNR) to a recommendable value of 40 decibels and above (> 40dB). In other words, the reversible test establishes that there is not much observable difference between the reconstructed cover image (C') and the original cover image (C). Therefore, the higher the PSNR, the better.

5. Result and Findings

The outcome of the ensembled model, which is the Hybrid Deep Learning Model (HDLM), was obtained by using payload capacity, Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) as the metrics. Also, the security test was carried out using the stegoExpose.

These results were obtained and compared with the existing models trained using a similar COCO image dataset. These results are tabulated in Table 1 and Table 2.

5.1.Evaluation Using Payload Capacity (bpp)

Table 1. HDLM against other Deep learning models based on payload capacity

Models	Payload Capacity (bpp)
SteganoGAN [11]	4.4
Encoder-Decoder: DCGAN [22]	24
FNNSteg [23]	4
HiddingGAN [12]	4
HDLM (our model)	24.83

5.2. *Evaluation Using Payload Capacity (bpp)*

Table 2. HDLM against other Deep learning models based on PSNR and SSIM

Models	PSNR (dB)	SSIM
SteganoGAN [11]	36.33	0.88
Encoder-Decoder: DCGAN [22]	34.55	0.95
HiddingGAN [12]	33.16	0.96
HDLM (our model)	41.44	0.97

Experiments in this study demonstrated that, compared to other algorithms, the extraction impact of the HDLM method is quite good. However, transferring secret images is not the same as transferring secret textual information; when extracting the secret image, there is usually a relatively small loss, such as noise or changes in the pixel value of some pixels, which usually affects the overall image understanding after reconstruction or decoding. The HDLM algorithm has a larger payload capacity and PSNR value when compared to other algorithms. The proposed HDLM's SSIM value is satisfactory but not as strong as the SSIM values of some of the compared algorithms. Tables 1 and 2 show that, while the proposed HDLM's payload capacity (bpp), PSNR, and SSIM are good, the difference between the secret image and the reconstructed secret image is extremely minimal, and the influence on the original cover image is also small.

5.3. Security Test Using StegoExpose

To ensure the HDLM steganography system's security, we exposed the stego-images obtained by our HDLM to stegoExpose. This current steganalysis tool has been extensively used for identifying suspicious images. StegExpose is a steganography detection program that detects steganography in images.

In this study, a total number of 1,000 stego-images were generated. To test the system's security using stegoExpose, a random sample size was selected using a simple random technique, Yamane's sample size method, as shown in equation 3.

$$n = N/(1 + Ne^2) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 3$$

Where n = sampled stego-image size =?
N = total number of stego-images = 1000
e = level of precision. This value is set to be 0.05 because we are not certain of the level of variability in the total number of the stego-images.

Yamane's equation was adopted because the generated stego-image is small compared to the dataset used for the models' training, and a 285 sample size was obtained.

Figure 3 shows that the generated stego-images are stored in the same folder directory with the stegoExpose program. The stegoExpose receives the folder directory as input via the windows command, the obtained StegoExpose report is a comma-separated value file (.csv), and it is shown in Table 3

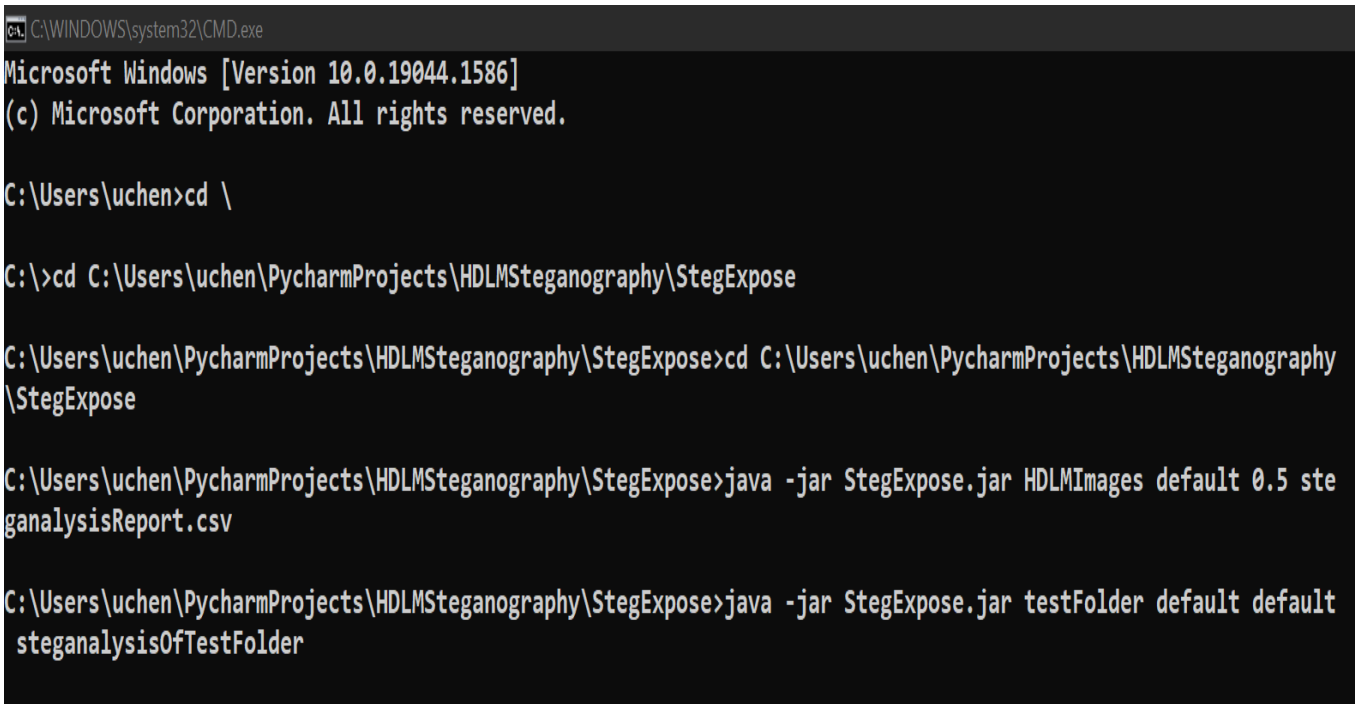

Fig. 3 StegoExpose Command Line Execution Interface

Table 3. HDLM StegoExpose Report

SN	Filename	Above the stego threshold?	Chi-Square	RS analysis	Fusion (mean)
1	stego (1).jpg	FALSE	0.01095085	0.063432081	0.034295218
2	stego (10).jpg	FALSE	0.010567242	0.016403342	0.013266601
3	stego (100).jpg	FALSE	0.008832068	0.029750216	0.018365141
4	stego (101).jpg	FALSE	0.004964846	0.005577701	0.055271273
5	stego (102).jpg	FALSE	0.03161545	0.091617445	0.060856052
6	stego (103).jpg	FALSE	0.039400657	0.029932385	0.027015404
7	stego (104).jpg	FALSE	0.017888555	0.078845258	0.04163583
8	stego (105).jpg	FALSE	0.012930021	0.050752425	0.061924625
9	stego (106).jpg	FALSE	0.098492958	0.01027998	0.012115045
10	stego (107).jpg	FALSE	0.070132976	0.024123695	0.042297043
11	stego (108).jpg	FALSE	0.012789889	0.008974523	0.009035064
12	stego (109).jpg	FALSE	0.084165415	0.052281507	0.049641518
13	stego (11).jpg	FALSE	0.023521177	0.026630297	0.093940419
14	stego (110).jpg	FALSE	0.010452686	0.013398235	0.054482027
15	stego (111).jpg	FALSE	0.002977124	0.004445806	0.006265595
16	stego (112).png	TRUE	0.483475599	1	0.741737799
17	stego (113).png	FALSE	0.035707498	0.028445343	0.064363293
18	stego (114).png	FALSE	0.08218266	0.06107636	0.07162951
19	stego (115).png	FALSE	0.031096265	0.070086075	0.00059117
20	stego (116).png	TRUE	0.577244389	1	0.537306782
21	stego (117).png	FALSE	0.035707498	0.028445343	0.064363293
22	stego (118).png	TRUE	0.919967738	1	0.798610891
23	stego (119).png	TRUE	0.315875764	1	0.641137797
24	stego (12).jpg	FALSE	0.057869592	0.021496516	0.02878275
.					
.					
.					
277	stego (91).jpg	FALSE	0.212747029	0.006440456	0.077701538
278	stego (92).jpg	FALSE	0.103234751	0.025145241	0.152016432
279	stego (93).jpg	FALSE	0.018294123	0.046278227	0.04047339
280	stego (94).jpg	FALSE	0.002532957	0.104950546	0.079863462
281	stego (95).jpg	FALSE	0.109536941	0.060565225	0.052784487
282	stego (96).jpg	FALSE	0.054521106	0.018637141	0.024264778
283	stego (97).jpg	FALSE	0.025981816	0.24911767	0.179714519
284	stego (98).jpg	FALSE	0.027681595	0.043539984	0.034832374
285	stego (99).jpg	FALSE	0.094002642	0.033570742	0.126269518

Table 3 shows an extract from the outcome of the StegoExpose on the HDLM steganography system. 285 stego-image samples are selected and saved in a folder directory fed into the StegoExpose steganalysis tool. It was discovered that the stegoExpose failed to detect 281 stego-images as suspicious and marked them as 'False' with rescaled range analysis values (RS Analysis Value) <1, but succeeded in detecting only four (4) stego-images as suspicious and marked them as 'True' with RS Analysis values = 1. The Chi-Square, Rescaled Range (RS) analysis, and the fusion (mean) values were calculated and used by the stegoExpose as parameters. Figures 4 and 5 depict the representations of Table 3 using charts.

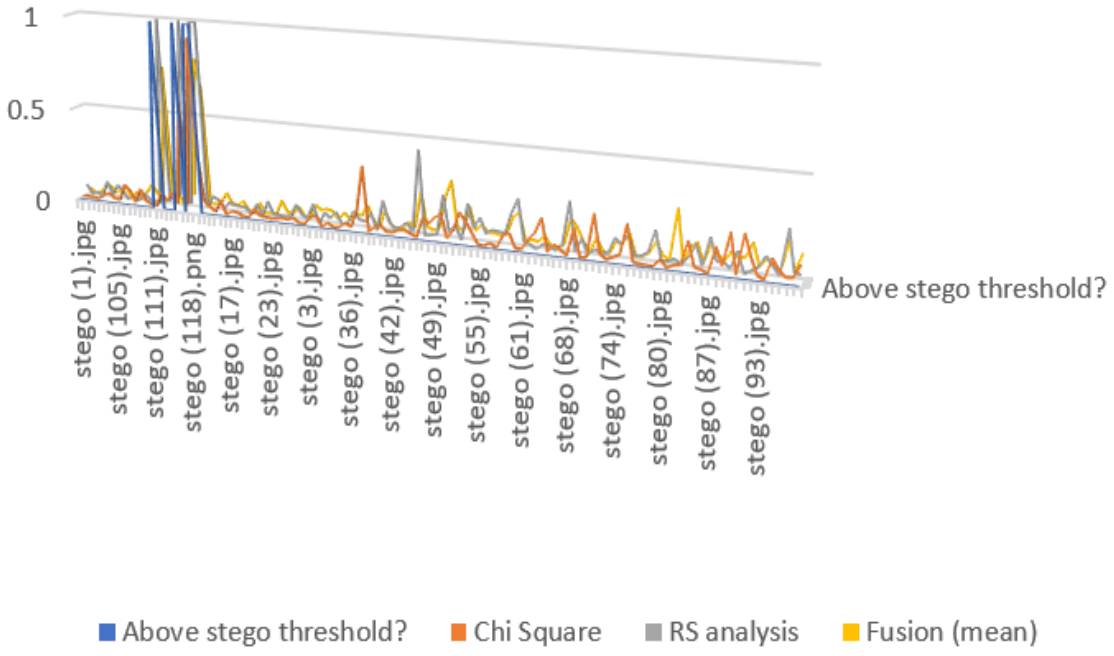

Fig. 4 StegoExpose RS Analysis

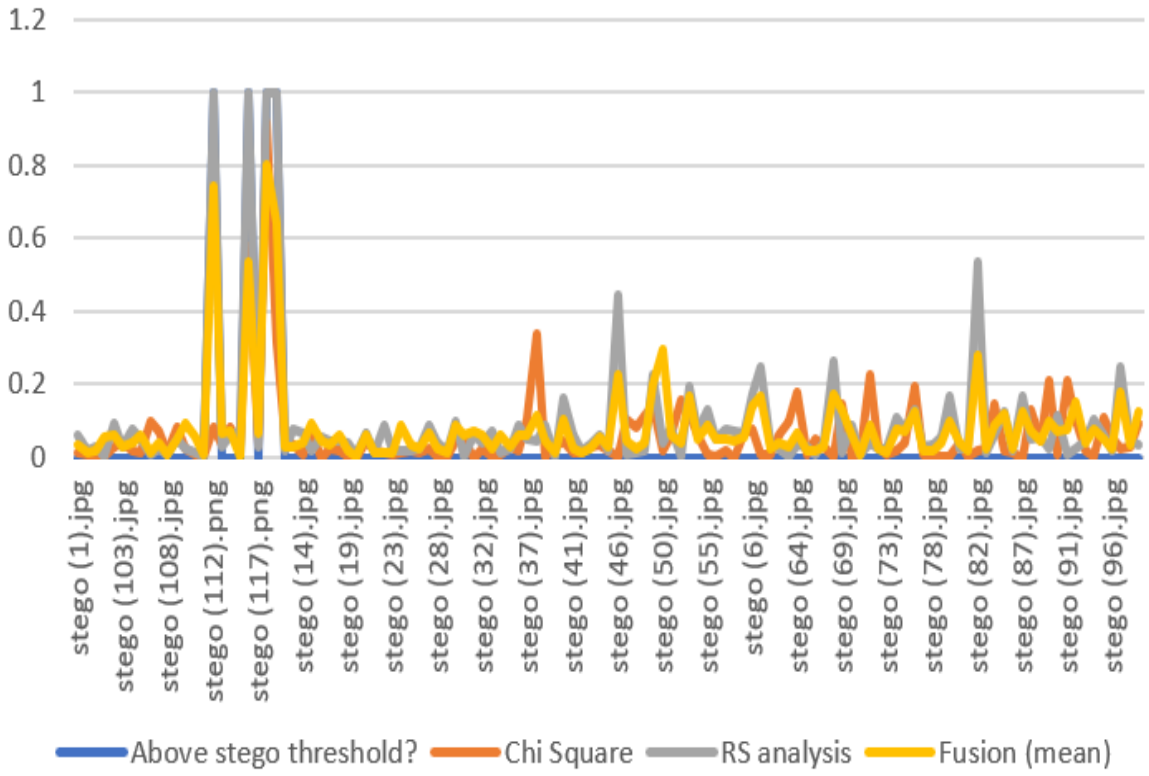

Fig. 5 Stego-image above 0.5 threshold value

6. Conclusion

The security of the HDLM Image steganography system was discovered to be very strong, as evident in Table 3 and Figures 4 and 5, where the threshold values for Chi-Square, RS analysis, and Fusion (mean) are all below 0.5, except for the four true detections of suspected stego-images out of the 285 selected samples of images fed into the StegoExpose steganalysis tool.

References

[1] ALabaichi A, Al-Dabbas M. A. A. A. K, & Salih A, Image Steganography Using the Least Significant Bit and Secret Map Techniques, International Journal of Electrical & Computer Engineering. 10(1) (2020) 2088-8708.

[2] Rachael O, Misra S, Ahuja R, Adewumi A, Ayeni F, & Mmaskeliunas R, Image Steganography and Steganalysis Based on Least Significant Bit (LSB), In Proceedings of ICETIT, Springer, Cham. (2020) 1100-1111.

[3] Yari I. A, & Zargari S, An Overview and Computer Forensic Challenges in Image Steganography, In 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom), and IEEE Smart Data (SmartData), IEEE. (2017) 360-364.

[4] Alsaidi N, Alshareef M, Alsulami A, Alsafri M, & Aljahdali A, Digital Steganography in Computer Forensics. Int. J. Comput. Sci. Inf. Secur. 18(5) (2020) 54-61.

[5] Laishram D, & Tuithung T, A Survey on Digital Image Steganography: Current Trends and Challenges, In Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT). (2018) 26-27.

[6] Idakwo M. A, Muazu M. B, Adedokun E. A, & Sadiq B. O, An Extensive Survey of Digital Image Steganography: State of the Art. ATBU Journal of Science, Technology, and Education. 8(2) (2020) 40-54.

[7] Verma V, Muttoo S. K, & Singh V. B, Enhanced Payload and Trade-Off for Image Steganography Via a Novel Pixel Digits Alteration, Multimedia Tools and Applications. 79(11) (2020) 7471-7490.

[8] Öfverstedt J, Lindblad J, & Sladoje N, Stochastic Distance Transform: Theory, Algorithms, and Applications, Journal of Mathematical Imaging and Vision. 62(5) (2020) 751-769.

[9] Yi Y, Hui L, Minghui X, & Yuntao W, High-Resolution Light Field Display Simulation Based on Frequency Domain Translation, Laser & Optoelectronics Progress. 59(1) (2020) 0107001.

[10] Duan X, Liu N, Gou M, Wang W, & Qin C, SteganoCNN: Image Steganography with Generalization Ability Based on Convolutional Neural Network, Entropy. 22(10) (2020) 1140.

[11] Zhang K. A, Cuesta-Infante A, Xu L, & Veeramachaneni K, SteganoGAN: High Capacity Image Steganography with GANs, arXiv preprint arXiv:1901.03892. (2019).

[12] Wang Z, Gao N, Wang X, Xiang J, Zha D, & Li L, HidingGAN: High Capacity Information Hiding with the Generative Adversarial Network, In Computer Graphics Forum. 38(7) (2019) 393-401.

[13] Yedroudj M, Chaumont M, Comby F, Oulad Amara A, & Bas P, Pixels-off: Data-Augmentation Complementary Solution for Deep-Learning Steganalysis, In Proceedings of the 2020 ACM Workshop on Information Hiding and Multimedia Security. (2020) 39-48.

[14] Ruan F, Zhang X, Zhu D, Xu Z, Wan S, & Qi L, Deep Learning for Real-Time Image Steganalysis: A Survey, Journal of Real-Time Image Processing. 17(1) (2020) 149–160. https://doi.org/10.1007/s11554-019-00915-5

[15] Kodovský J, Fridrich J, & Holub V, Ensemble Classifiers for Steganalysis of Digital Media, IEEE Transactions on Information Forensics and Security. 7 (2012) 432-444.

[16] Babu J, Rangu S, & Manogna P, A Survey on Different Feature Extraction and Classification Techniques Used in Image Steganalysis, Journal of Information Security. 8(3) (2017) 186-202. https://doi.org/10.4236/jis.2017.83013

[17] Boehm B, Steg Expose - A Tool for Detecting LSB Steganography. (2014) 1–11. http://arxiv.org/abs/1410.6656

[18] Zhou X, Yao C, Wen H, Wang Y, Zhou S, He W, & Liang J, East: An Efficient and Accurate Scene Text Detector, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2017) 5551-5560.

[19] Voravuthikunchai W, Crémilleux B, & Jurie F, Histograms of Pattern Sets for Image Classification and Object Recognition, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2014) 224-231.

[20] Arjovsky M, Chintala S, & Bottou L, Wasserstein Generative Adversarial Networks, Proceedings of the 34th International Conference on Machine Learning, in Proceedings of Machine Learning Research. 70 (2017) 214-223. Available: https://proceedings.mlr.press/v70/arjovsky17a.html.

[21] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, & Bengio Y, Generative Adversarial Networks, Communications of the ACM. 63(11) (2020) 139-144.

[22] Subramanian N, Cheheb I, Elharrouss O, Al-Maadeed S, & Bouridane A, End-to-End Image Steganography Using Deep Convolutional Autoencoders, IEEE Access. 9 (2021) 135585-135593.

[23] Kishore V, Chen X, Wang Y, Li B, & Weinberger K. Q, Fixed Neural Network Steganography: Train the Images, Not the Network, In International Conference on Learning Representations. (2021).

[24] Mittal S, A Survey on Modeling and Improving the Reliability of DNN Algorithms and Accelerators, Journal of Systems Architecture. 104 (2020) 101689.

[25] Harms J, Lei Y, Wang T, Zhang R, Zhou J, Tang X, & Yang X, The Paired Cycle GAN Based Image Correction for Quantitative Cone Beam Computed Tomography, Medical Physics. 46(9) (2019) 3998-4009.