IMAGE STEGANOGRAPHY USING CNN

Shourya Chambial[1], Dhruv Sood[2]

[1]School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, 632014, India
[2]School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, 632014, India

---***---

Abstract - *The unfamiliarity and curiosity in the subject were the key reasons for choosing steganography from the list of available project ideas. Another argument is that it ensures data security. Even if a hacker (or intruder) has access to our image files, they will be unable to access our information; that is, even if the intruder manages to hack and gain access to our system, they will be unable to access our data. This feature adds a lot to the appeal. The project is about utilising images to learn about text stenography. For images, image steganography is used, and the relevant data is decrypted in order to get the message image. Image steganography is researched, and one of the methods is used to demonstrate it, because it may be done in a variety of ways. Steganography is the practise of concealing data, such as text, photos, or audio files, within another image or video file. In a word, steganography's fundamental goal is to conceal the desired information within any image, audio, or video that does not appear to be hidden just by looking at it. Image-based Steganography is based on a basic concept. Images are made up of digital data (pixels) that define what's inside the image, which is usually the colours of all the pixels. Since we know that every image is made up of pixels, each of which has three values (red, green, blue). We anticipate that the image we want to conceal will be totally absorbed by the cover image, making it invisible to intruders.*

***Key Words*: Image Steganography, Steganography, Steganography using CNN, Deep learning, GAN, LSB**

1.INTRODUCTION

A technique used for hiding secret data within a standard, public file or message to avoid detection is known as Steganography. The secret data is then extracted once it arrives at its destination. Steganography, in conjunction with encryption, can be used to further conceal or safeguard data.

Using deep convolutional neural networks, a full-sized colour image is concealed inside another image (called Cover image) with minimum changes in appearance. The hidden image will then be revealed by combining a "reveal" network with the created image.

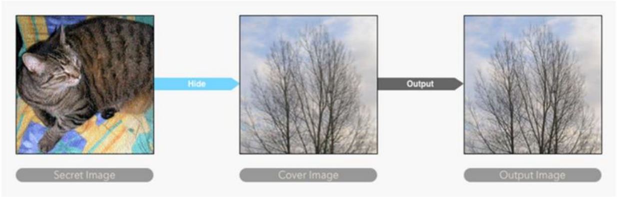

Fig -1: Overview

The first image in the above example is the secret image that must be kept concealed, while the second image is the cover image that will be used to conceal the secret image. The third image, which is essentially a repainted cover image, is the result of the hiding procedure. Some of the most common steganographic approaches involve the manipulation of the least significant bits (LSB) of the images to place the secret information, which can be achieved adaptively or even uniformly through simple replacement or through methods or through methods that are more advanced. In our project we will be using convolutional neural networks (CNN) for hiding an image inside another image. The advantage is a more efficient steganography.

There are disadvantages to this approach also, as on enhancing the image by a big factor, one can find out the original image but it is very hard to do this and to completely reconstruct the original image from the output image. We have used the TINY IMAGENET DATABASE.

2. Related work

2.1 Summary of various methods

Following a review of all available frameworks, the methodologies are primarily divided into three categories: traditional image steganography methods, CNN-based image steganography methods, and GAN-based image steganography methods. Traditional methods are frameworks that employ methods unrelated to machine learning or deep learning algorithms. The LSB technique is used in many traditional methods. CNN-based methods use deep convolutional neural networks to embed and extract secret messages, whereas GAN-based methods use some GAN variants.

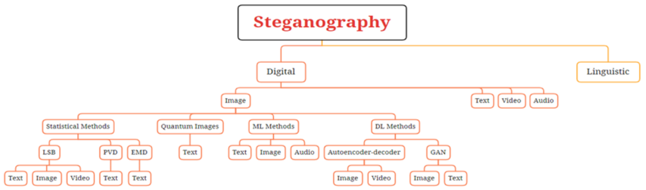

Fig -2: Classification of current methods on basis of the method and media use

2.2 Review on various schemes

Image steganography is the process of concealing information in a cover image, which might be text, image, or video. The secret information is concealed in such a way that it cannot be seen by human sight. Deep learning technology, which has recently gained popularity as a strong tool in a variety of applications, including image steganography, has gotten a lot of attention. The primary goal of this study is to analyze and give an explanation for the numerous deep studying algorithms to be had inside the field of photograph steganography which we came across in our referred papers. Traditional methods, Convolutional Neural Network-based methods, and General Adversarial Network-based methods are the three types of deep learning approaches utilized for image steganography. This study includes a detailed explanation of the datasets used, experimental set-ups explored, and commonly used assessment measures, in addition to the methodology. For simple reference, a table summarizing all of the details is also supplied.

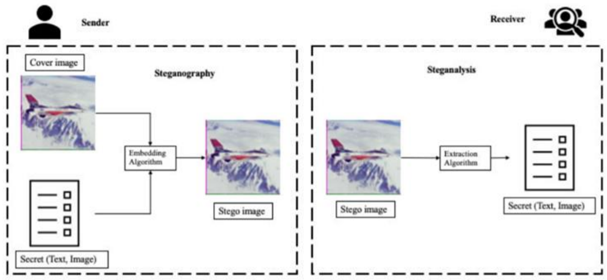

Fig -3: General principle of steganography

2.3 Steganography techniques that are based on traditional methods

Image steganography is traditionally done using the Least Significant Bits (LSB) substitution method. The pixel quality of images is normally higher, but not all of the pixels are used. The LSB approach is based on the notion that changing a few pixel values will not result in noticeable changes. The secret data is transformed into a binary format. The least significant bits in the noisy area are determined by scanning the cover image. The LSBs of the cover picture are then replaced with the binary bits from the secret image. The substitute approach must be used with caution, as overloading the cover picture may result in noticeable alterations that reveal the secret information's presence.

Table -1: Traditional Methods

Dataset	Metrics	Advantages	Disadvantages
Lena	PSNR	Time to compute is reduced. The image is a coded message in and of itself.	It is insecure.
Lena and Baboon	PSNR and MSE	Time to compute is reduced. The image is a coded message in and of itself. Any image format is acceptable.	When compared to deep learning approaches, security is a concern.
1 RGB Image	PSNR and Time	Time to compute is reduced. It embeds data with ease. Steganography and steganalysis are not reliant on one other.	It is insecure. Text is used to communicate secret information.

2.4 GAN–BASED methods used in Steganography

General Adversarial Networks (GAN) are a type of deep CNN. For image generation tasks, a GAN employs game theory to train a generative model with an adversarial process. In GAN architecture, two networks – generator and discriminator networks – compete to generate a perfect image. The data is fed into the generator model, and the output is a close approximation of the given input image. The discriminator networks classify the generated images as either false or true. The two networks are trained in such a way that the generator model attempts to imitate the input data as closely as possible with as little noise as possible.

Image steganography existing methods using a GAN architecture can be categorized into five types: a three-network-based GAN model, cycle-GAN-based architectures, sender-receiver GAN architectures, coverless models in which the cover image is generated randomly rather than being given as input, and an Alice, Bob, and Eve based model.

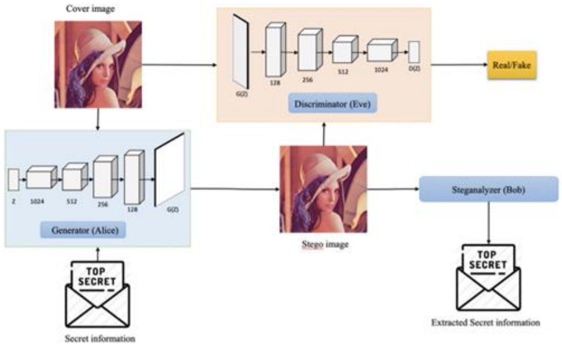

Fig -4: GAN Overview

A GAN model is made up of two main components: the generator and the discriminator. In some of the methods for image steganography, a new network called the steganalyzer is introduced. These three components' primary functions are as follows:

- A model, G, for generating stego images from the cover image and the random message.
- A discriminator model, D, to determine whether the generated image from the generator is real or fake.
- A steganalyzer, S, to determine whether or not the input image contains confidential secret data.

Table -2: GAN Based Methods

Architecture	**Dataset**	**Advantages**	**Disadvantage**
Alice, Bob, Eve	BOSSbase and celebA	Embedding does not necessitate domain knowledge.	Messages are used instead of images. Grid search for embedding scheme selection consumes time.
DCGAN	celebA and BOSSbase	Game-theoretic formulation is used	Steganalyzer is used to provide the probability by itself, which lead in additional computational cost and overhead.
GAN	CelebA	Unlimited number of cover images are be generated	Generator and discriminator are shared. Images that have been generated are not natural.
DCGAN	CelebA	Image creation that is more realistic Extremely safe	It provides probabilistic information rather than secret information. There is no information on how to obtain the secret data.

2.5 Steganography techniques that are based on CNN

The encoder-decoder architecture is greatly influenced by steganography employing CNN models. The encoder receives two inputs: the cover picture and the secret image, which are used to construct the stego image, and the decoder receives the stego image, which outputs the embedded secret image. The core principle remains the same, however different techniques have experimented with various structures. Different algorithms change in how the input cover picture and the secret image are concatenated, while variations in the convolutional layer and pooling layer are to be expected. The number of filters used, the strides taken, the filter size utilised, the activation function used, and the loss function used differ from one approach to the next. One thing to keep in mind is that the cover image and the secret image must be the same size, so that every pixel of the secret image is dispersed throughout the cover image.

Convolution is a type of linear operation that expresses the degree of overlap between two functions when they are shifted over each other. Convolutional networks are simple neural networks with at least one layer that uses convolution instead of ordinary matrix multiplication. The following are some of the benefits of using a CNN-based architecture for encoding and decoding:

- CNN extracts visual features automatically.
- CNN essentially down samples the image using nearby pixel information, first by convolution, and then by using a prediction layer at the end.
- CNN is more accurate and performs effectively.

One can get a decent idea of the patterns of natural images by using a deep neural network, in this case a CNN. The network will be able to determine which areas are redundant, allowing additional pixels to be buried in certain areas. The amount of hidden data can be raised by saving space on superfluous areas. Because the structure and weights can be randomized, the network will conceal data that is inaccessible to anyone who does not have the weights.

Table -3: CNN Based Methods

Architecture	**Dataset**	**Advantages**	**Disadvantages**
Encoder-decoder with SCR	ImageNet	Highly secure and dependable	The loss used isn't ideal. In black or white areas, visible noise might be seen.
Encoder-decoder with VGG-base	COCO and wikiart.org	It is not necessary to have prior domain knowledge. Since the created image is unrelated to the secret information, it is extremely secure.	High number of images are required in computation purpose.
CNN	ImageNet and Holiday	The image is a coded message in and of itself. The simplest and most basic architecture is chosen. To speed up training, a new error back propagation function is implemented.	However, the image is only 64x64 pixels in size, which is quite little. The images in the input are simply concatenated.
U-Net	ImageNet	The image is a coded message in and of itself. The simplest and most basic architecture is chosen.	However, the image is only 64x64 pixels in size, which is quite little. The images in the input are simply concatenated.
Encoder-decoder	ImageNet	The image is a coded message in and of itself.	However, the image is only 64x64 pixels in size, which is quite little.

3. Background

3.1 General Architecture on Topic chosen

The entire system consists of 3 CNNs:

1. The Preparation Network: This network prepares the secret image to be hidden the purpose is to transform the color-based pixels to more useful features so that in can be encoded such as edges. It contains 50 filters of (3X3, 4X4, 5X5) patches there are total 6 layers of this kind.

2. The Hiding Network: This network will take the output of the preparation network and will then create a Container Image. It is a CNN with 5 convolutional layers that have 50 filters of (3X3, 4X4, 5X5) patches. This layer consists of total of 15 convolutional layers in this network.

3. The Reveal Network: Converts the Container image to the original image this network is used for decoding. Which has 5 convolutional layers that have 50 filters of (3X3, 4X4, 5X5) patches. This layer consists of total of 15 convolutional layers in this network.

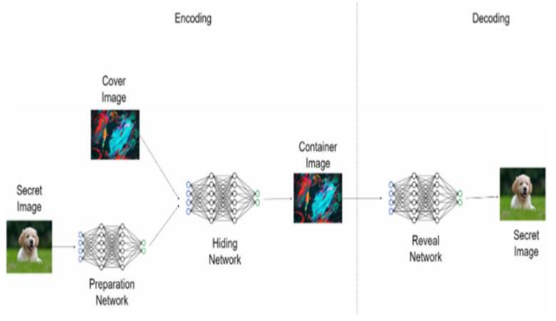

Fig -5: Encoder and Decoder Overview

Some small amount of noise is also added to the images that are the output of the second network in order to make sure that the system does not simply encode the secret image in the LSB.

We want to do image steganography, which involves hiding an image inside one cover image. The concealed images included in the cover image must be retrievable with minimal loss. The encoded cover image must resemble the original cover image in appearance.

We transfer secret images over the prep network, then concatenate the resulting data with the carrier image before sending it over the Hiding network. The idea of having decoders, one for each secret picture, to obtain the secret image from the container image is then implemented.

To make our image retrieval model more secure, we expand on the concept of putting instead, a secret image with noise in the original cover image placing the secret photos on the original cover's LSBs image.

Following is a brief overview of the encoder/decoder framework used in our technique:

ENCODER: A prep network that corresponds to secret image input. The outputs of the prep network are combined with the cover picture and passed via the Hiding network.

DECODER: The decoder network is made up of reveal networks, each of which has been taught to decode its own message.

The following is the typical framework:

Prep Networks: Each prep network is made up of two layers stacked on top of each other. Each layer is made up of three independent Conv2D layers. These three Conv2D layers have 50, 10, and 5 channels, respectively, with kernel sizes of 3, 4, and 5 for each layer. Along both axes, the stride length remains constant at one. To preserve the output image in the same dimensions, appropriate padding is provided to each Conv2D layer. After each Conv2d layer, a Relu activation is applied.

Concealing Network: The concealing network is a three-layer aggregation. Each of these layers is made up of three individual Conv2D layers. The Conv2D layers in the hidden network have a similar basic structure to the Conv2D levels in the Prep Network.

Reveal Network: The reveal network has a similar basic design to the hidden network, with three levels of Conv2D layers that are comparable in shape.

4. Proposed algorithm

The system that we have implemented in this project uses a CNN to hide an image inside another image hence making the secrete image effectively invisible to the observer. The neural networks that we have used in our work determines where in the cover image it can hide the information and hence, is a more efficient process than LSB manipulation. We have also trained a decoder network to reveal the secret image from the container image. This process is done in such a way that the container image does not change much and the changes to the container image are not discernible.

It can be safely assumed that the intruder does not have access to the original image. The secret image is effectively hidden in all 3 colour channels of the container image.

3 different activation functions were used to test the network along with various learning rates.

The activation functions used were:

1. Rectified Linear Unit (RELU) {de-facto for most CNN networks}

2. Tanh

3. Scaled Exponential Linear Unit (SELU) {has been used to avoid the vanishing gradient problem caused by RELU activation function}

Advantages: -

1. CNNs learn the most efficient way to hide the secret image inside the container image which is unpredictable for us. As a result, it adds to the security of steganography by making the process unpredictable.

2. The container image is note altered a lot which makes the changes harder to spot.

3. The changes can only be spotted by the decoder network that is specifically trained to do it.

4. The process is flexible that means that before feeding to network, some changes can be made to the image to make the process more secure.

Disadvantages: -

1. This is not a full proof method and the changes can be spotted by a specially trained network and the original image can be somewhat retrieved.

We have chosen this method because the advantages clearly outweigh the disadvantages.

4.1 Our dataset

We will be using the Tiny ImageNet which contains 100000 images divided into 200 classes (500 images per class) and downsized to 64 x 64 coloured images. It consist of 500 training images, 50 validation images, and 50 test images in each class.

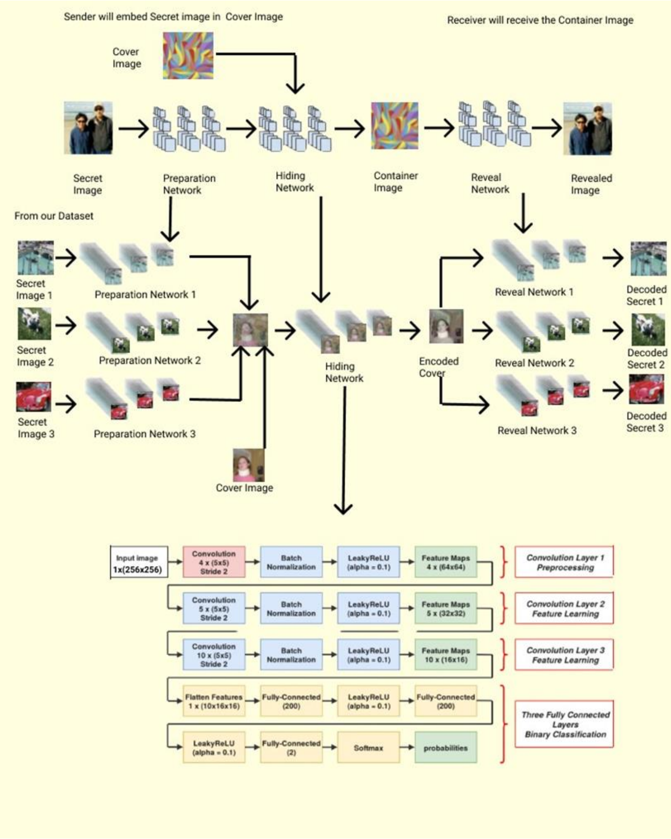

Fig -6: Architecture Diagram

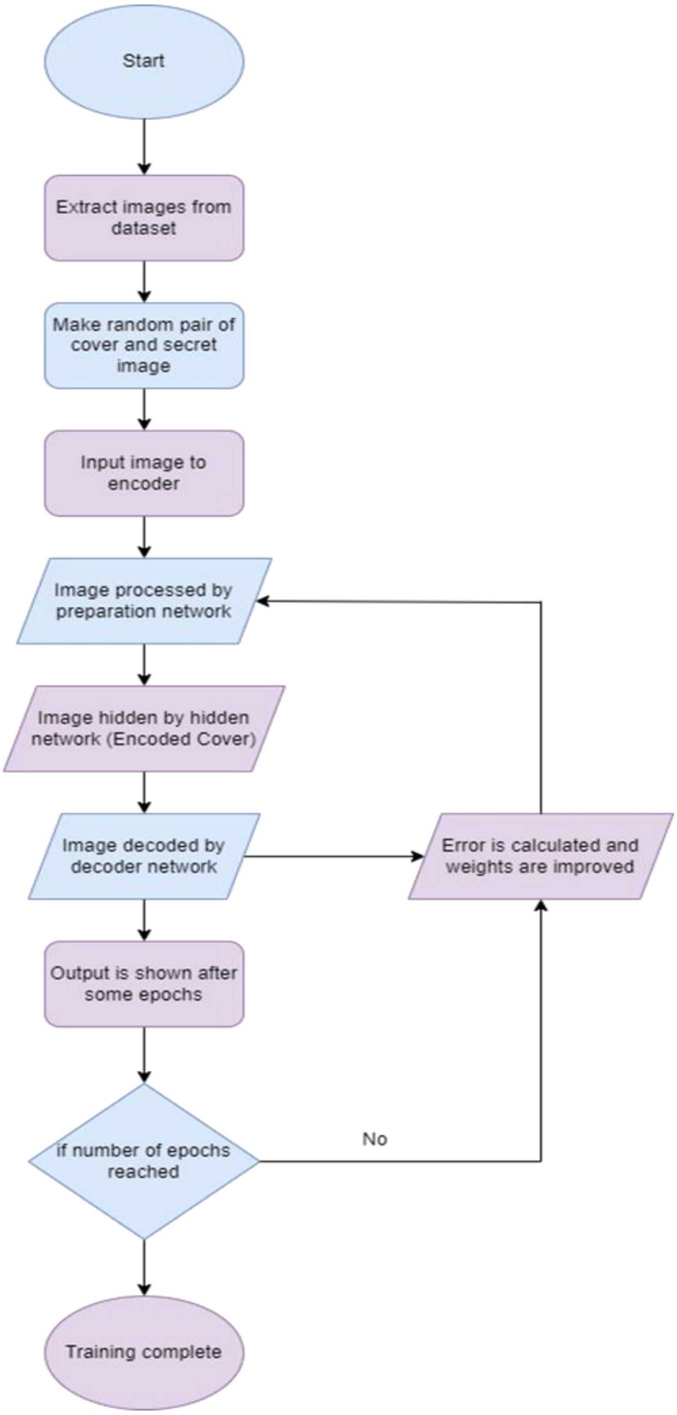
Start
Extract images from dataset
Make random pair of cover and secret image
Input image to encoder
Image processed by preparation network
Image hidden by hidden network (Encoded Cover)
Image decoded by decoder network
Error is calculated and weights are improved
Output is shown after some epochs
if number of epochs reached
No
Training complete

Fig -7: Flow Chart

5. Results

ImageNet is a massive dataset that contains images from the WordNet hierarchy, with each node holding hundreds of images. The image on ImageNet has no copyrights and just contains links or thumbnails to the original image. Images of various sizes make up the dataset. The number of photographs, the classes they belong to, the background, and the image size can all be customized based on the requirements.

ImageNet's Tiny ImageNet dataset is a subset of the larger ImageNet dataset. The dataset consists of 100,000 photos divided into 200 classes (500 images per class) that have been reduced to 6464 coloured images. There are 500 training images, 50 validation images, and 50 test photographs in each class.

The quality of image steganography is assessed using a variety of approaches. Each of these methods evaluates a different component of the steganographic outcome. Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), and Structured Similarity Index Measure (SSIM) are some of the well-known approaches.

Mean Square Error (MSE):-

The averaged value of the square of the pixel-by-pixel difference between the original image and the stego-image is the Mean Square Error. It provides a measure of the cover image error caused by the data embedding procedure. A lower MSE value denotes a high-quality embedding.

Peak Signal to Noise Ratio (PSNR):-

PSNR is another useful metric for determining the degree of embedding-induced distortion in the cover image. It's the ratio of a signal's greatest possible value to the power of distortion noise (MSE). It is expressed in decibels (dB). A higher PSNR value denotes a higher-quality embedding.

Structured Similarity Index Measure (SSIM):-

SSIM is a comparison statistic used to see how similar the cover picture and the stego-image are. The perceived difference between the two images is measured. We were able to achieve MSE value of 0.006259255611669444, PSNR value of 76.21412244818211 dB, and SSIM value of 0.96512678174912

In this section the results of the proposed methodology are compared with the existing.

Table -3: Comparative Study

Parameter	[31]	[32]	[33]	Proposed Algorithm
Dataset	Lena and Baboon	ImageNet	1 RGB image	Tiny ImageNet
Method used	GAN	GAN	LSB	CNN
Architecture	Traditional Method	U NET	Traditional Method	Encoding and decoding
MSE	0.13	-	-	0.006
PSNR	56.95	40.66	62.53	76.214
SSIM	-	0.964	-	0.965

6. Conclusion and future work

- The network was constructed and is performing well. Even though it has shown good performance encoding and decoding, it is not a full proof system as every technology has its shortcomings.
- As stated in the disadvantages, there is a lot of scope of improvement in this project as there is a chance that networks can be trained specifically for detecting that an image has been hidden inside a given image.
- This project can be improved upon by creating a better architecture or by altering the architecture of this network and improving its performance further, making the secret image tougher to decode by any program other than the decoder and to make it more difficult to detect the presence of any secret image inside the encoded image.

REFERENCES

[1] H. Kato, K. Osuge, S. Haruta and I. Sasase, "A Preprocessing by Using Multiple Steganography for Intentional Image Downsampling on CNN-Based Steganalysis," in IEEE Access, vol. 8, pp. 195578-195593, 2020, doi: 10.1109/ACCESS.2020.3033814.

[2] Jin, Z., Yang, Y., Chen, Y. and Chen, Y., 2020. IAS-CNN: Image adaptive steganalysis via convolutional neural network combined with selection channel. International Journal of Distributed Sensor Networks, 16(3), p.1550147720911002.

[3] Xiang, Z., Sang, J., Zhang, Q., Cai, B., Xia, X. and Wu, W., 2020. A new convolutional neural network-based steganalysis method for content-adaptive image steganography in the spatial domain. IEEE Access, 8, pp.47013-47020.

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

[5] Duan, X., Guo, D., Liu, N., Li, B., Gou, M. and Qin, C., 2020. A new high capacity image steganography method combined with image elliptic curve cryptography and deep neural network. IEEE Access, 8, pp.25777-25788.

[6] Li, Q., Wang, X., Wang, X., Ma, B., Wang, C., Xian, Y. and Shi, Y., 2020. A novel grayscale image steganography scheme based on chaos encryption and generative adversarial networks. IEEE Access, 8, pp.168166-168176.

[7] Tang, W., Li, B., Tan, S., Barni, M. and Huang, J., 2019. CNN-based adversarial embedding for image steganography. IEEE Transactions on Information Forensics and Security, 14(8), pp.2074-2087.

[8] Yu, X., Tan, H., Liang, H., Li, C.T. and Liao, G., 2018, December. A multi-task learning CNN for image steganalysis. In 2018 IEEE International Workshop on information forensics and security (WIFS) (pp. 1-7). IEEE.

[9] Yuan, Y., Lu, W., Feng, B. and Weng, J., 2017, June. Steganalysis with CNN using multi-channels filtered residuals. In International Conference on Cloud Computing and Security (pp. 110-120). Springer, Cham.

[10] Duan, X., Guo, D., Liu, N., Li, B., Gou, M. and Qin, C., 2020. A new high-capacity image steganography method combined with image elliptic curve cryptography and deep neural network. IEEE Access, 8, pp.25777-25788.

[11] Pibre, L., Pasquet, J., Ienco, D. and Chaumont, M., 2020. Deep learning is a good steganalysis tool when embedding key is reused for different images, even if there is a cover source mismatch. Electronic Imaging, 2016(8), pp.1-11.

[12] Kim, J., Park, H. and Park, J.I., 2020. CNN-based image steganalysis using additional data embedding. Multimedia Tools and Applications, 79(1), pp.1355-1372.

[13] Zou, Y., Zhang, G. and Liu, L., 2019. Research on image steganography analysis based on deep learning. Journal of Visual Communication and Image Representation, 60, pp.266-275.

[14] Kweon, H., Park, J., Woo, S. and Cho, D., 2021. Deep Multi-Image Steganography with Private Keys. Electronics, 10(16), p.1906.

[15] You, W., Zhang, H. and Zhao, X., 2020. A Siamese CNN for image steganalysis. IEEE Transactions on Information Forensics and Security, 16, pp.291-306.

[16] Zhang, C., Lin, C., Benz, P., Chen, K., Zhang, W., & Kweon, I. S. (2021). A brief survey on deep learning based data hiding, steganography and watermarking. arXiv preprint arXiv:2103.01607.

[17] SUBRAMANIAN, N. (2021). Image Steganography Using Deep Learning Methods to Detect Covert Communication in Untrusted Channels (Master's thesis).

[18] Lu, S. P., Wang, R., Zhong, T., & Rosin, P. L. (2021). Large-Capacity Image Steganography Based on Invertible Neural Networks. In Proceedings of the IEEE/Computer Vision and Pattern Recognition (pp. 10816-10825).

[19] Ruiz, H., Chaumont, M., Yedroudj, M., Amara, A. O., Comby, F., & Subsol, G. (2021, January). Analysis of the scalability of a deep-learning network for steganography “Into the Wild”.Springer, Cham.

[20] Byrnes, O., La, W., Wang, H., Ma, C., Xue, M., & Wu, Q. (2021). Data Hiding with Deep Learning: A Survey Unifying Digital Watermarking and Steganography. arXiv preprint arXiv:2107.09287.

[21] Zhangjie, F., Enlu, L., Xu, C., Yongfeng, H., & Yuting, H. (2021). Recent Advances in Image Steganography Based on Deep Learning. Journal of Computer Research and Development, 58(3), 548.

[22] Chang, C. C., Wang, X., Chen, S., Echizen, I., Sanchez, V., & Li, C. T. (2021). Deep Learning for Reversible Steganography: Principles and Insights. arXiv preprint arXiv:2106.06924.

[23] Selvaraj, A., Ezhilarasan, A., Wellington, S. L. J., & Sam, A. R. (2021). Digital image steganalysis: A survey on paradigm shift from machine learning to deep learning based techniques. IET Image Processing.

[24] Chang, C. C. (2021). Neural Reversible Steganography with Long Short-Term Memory. Security and Communication Networks, 2021.

[25] Kweon, H., Park, J., Woo, S., & Cho, D. (2021). Deep Multi-Image Steganography with Private Keys. Electronics, 10(16), 2021.

[26] Bashir, B., & Selwal, A. (2021). Towards Deep Learning-Based Image Steganalysis: Practices and Open Research Issues. Available at SSRN 3883330.

[27] Salunkhe, S., & Bhosale, S. Feature Extraction Based Image Steganalysis Using Deep Learning.

[28] Das, A., Wahi, J. S., Anand, M., & Rana, Y. (2021). Multi-Image Steganography Using Deep Neural Networks. arXiv preprint arXiv:2101.00350.

[29] Cui, J., Zhang, P., Li, S., Zheng, L., Bao, C., Xia, J., & Li, X. (2021). Multitask Identity-Aware Image Steganography via Minimax Optimization. arXiv preprint arXiv:2107.05819.

[30] Cherian, R. E. Cryptography and Deep Neural Network: An Art of Hiding Data in Images.

[31] A. Arya and S. Soni, ''Performance evaluation of secrete image steganography techniques using least significant bit (LSB) method,'' Int. J. Comput. Sci. Trends Technol., vol. 6, no. 2, pp. 160–165, 2018.

[32] X. Duan, K. Jia, B. Li, D. Guo, E. Zhang, and C. Qin, ''Reversible image steganography scheme based on a U-Net structure,'' IEEE Access, vol. 7, pp. 9314–9323, 2019.

[33] K. A. Al-Afandy, O. S. Faragallah, A. Elmhalawy, E.-S.-M. El-Rabaie, and G. M. ElBanby, ''High security data hiding using image cropping and LSB least significant bit steganography,'' in Proc. 4th IEEE Int. Colloq. Inf. Sci. Technol. (CiSt), Oct. 2016, pp. 400–404.