www.ijbar.org
ISSN2249-3352(P)2278-0505(E)
CosmosImpactFactor-5.86

AN IMPROVED FILE SECURITY SYSTEM USING DATA HIDING TECHNIQUES

Dr. C.Madan Kumar[1*], Sathupelli Rachana[2], Merugu Sai[2], Kokkirala Deekshitha[2], Mothe Shivaram Reddy[2]

[1]Assistant Professor, [2]UG Student, [1,2]Department of Computer Sciences and Engineering,

[1,2]Vaagdevi College of Engineering (UGC - Autonomous), Bollikunta, Warangal, Telangana.

[*]Corresponding author: Dr. C.Madan Kumar

ABSTRACT

In an age of increasing digital communication and data transfer, ensuring the security and privacy of sensitive information is paramount. Steganography, the art of hiding information within other data, has been used for centuries. In the digital realm, it plays a critical role in secure communication and information concealment. Traditional steganography methods often involve embedding information within a single image. While effective, this approach may be susceptible to detection, as single-image steganography can leave detectable traces, especially under sophisticated analysis. The primary challenge is to develop a robust system for multiple image steganography that can securely hide sensitive files within a set of images. This involves designing algorithms that distribute the information effectively across the images while maintaining imperceptibility and ensuring reliable extraction. Therefore, the rise of cyber threats and privacy concerns, there's a growing need for advanced techniques to protect sensitive files from unauthorized access or interception. Multiple image steganography, an emerging field, offers the potential for heightened security by spreading information across multiple images, making it even more challenging for potential adversaries to detect or extract. The project seeks to enhance file security by leveraging advanced techniques in multiple image steganography. By distributing the information across a set of images, this research endeavors to develop a system capable of securely concealing sensitive files. The algorithms utilized in this approach are designed to ensure imperceptibility and robustness against detection efforts. This advancement holds great promise for significantly improving the security of file transmission and storage, safeguarding critical information from unauthorized access or interception.

Keywords: Multiple Image Steganography, Data Hiding Techniques, Steganographic Algorithms, Digital Communication Security, Secure File Transmission.

1.INTRODUCTION

In today's digital era, safeguarding sensitive information during communication and data transfer is of utmost importance. Throughout history, the practice of steganography, which involves concealing information within other data, has been utilized to achieve this goal. As we transition to digital mediums, steganography plays a crucial role in ensuring secure communication and information concealment.Traditionally, steganography focused on embedding information within a single image. While effective, this approach has its drawbacks, as it can be susceptible to detection, especially under sophisticated analysis. The challenge at hand is to develop a robust system for multiple image steganography that can securely hide sensitive files within a collection of images. This entails designing algorithms that efficiently distribute the information across the images while maintaining imperceptibility and ensuring reliable extraction.

The increasing prevalence of cyber threats and privacy concerns underscores the need for advanced techniques to protect sensitive files from unauthorized access or interception. Multiple image steganography, an emerging field, offers the potential for heightened security by dispersing information across multiple images. This makes it even more challenging for potential adversaries to detect or extract the concealed data.

This project aims to enhance file security by leveraging advanced techniques in multiple image steganography. The goal is to develop a system capable of securely concealing sensitive files by distributing information across a set of images. The algorithms employed in this approach are specifically designed to ensure imperceptibility and robustness against detection efforts. This advancement holds great promise for significantly improving the security of file transmission and storage, safeguarding critical information from unauthorized access or interception.

2. LITERATURE SURVEY

B. Sultan,et.al[1] In this work a multi-data deep learning steganography model has been developed using a well known deep learning model called Generative Adversarial Networks (GAN) more specifically using deep convolutional Generative Adversarial Networks (DCGAN). The model is capable of hiding two different messages, meant for two different receivers, inside a single cover image. The proposed model consists of four networks namely Generator, Steganalyzer Extractor1 and Extractor2 network. The Generator hides two secret messages inside one cover image which are extracted using two different extractors. The Steganalyzer network differentiates between the cover and stego images generated by the generator network. The experiment has been carried out on CelebA dataset. Two commonly used distortion metrics Peak signal-to-Noise ratio (PSNR) and Structural Similarity Index Metric (SSIM) are used for measuring the distortion in the stego image The results of experimentation show that the stego images generated have good imperceptibility and high extraction rates.

X. Liao,et.al[2] formulated adaptive payload distribution in multiple images steganography based on image texture features and provides the theoretical security analysis from the steganalyst's point of view. Two payload distribution strategies based on image texture complexity and distortion distribution are designed and discussed, respectively. The proposed strategies can be employed together with these state-of-the-art single image steganographic algorithms. The comparisons of the security performance against the modern universal pooled steganalysis are given. Furthermore, this article compares the per image detectability of these multiple images steganographic schemes against the modern single image steganalyzer. Extensive experimental results show that the proposed payload distribution strategies could obtain better security performance.

M. Srivastava,etal. [3] worked on two techniques for hiding information in the image. First, we do analysis on LSB for storing information bit. As the technique is known to all, the attacker will be able to easily reveal the information, this makes image steganography unsecured. Secondly, R-Color Channel encoding with RSA set of rules for offering extra protection to information in addition to our information hiding approach. The proposed approach makes use of a red color channel for hiding information bits and the following bits for RGB pixel values of the original image. This paper present the performance analysis of two most popular algorithms, LSB and RSA along with image steganography.

G. Benedict,etal. proposed a novel approach is proposed for slicing the secret data and storing it on multiple cover images. In addition, retrieval of this secret data from the cover images on the destination side has also been discussed. The data slicing ensures secure transmission of the vital data making it merely impossible for the intruder to decrypt the data without the encrypting details.

S. Mukhopadhyay,etal. [5] proposed a scheme for achieving steganography with multiple encrypted monochromatic images with keys obtained from a synchronized system of semiconductor lasers. The key selection scheme for steganography determines the robustness of the application. It is in this area that steganography may benefit from the properties of chaos synchronization. The encryption principle of the new algorithm is analyzed quantitatively by various statistical tests. The cover image used in the technique is also obtained from the visual representation of the chaotic sequences. This new scheme enjoys the benefit of added security, high key space, high embedding capacity, imperceptibility and robustness of the hidden information in conjunction with Least Significant Bit (LSB) based substitution. The result is important from the perspective of introducing a mechanism to multiplex and simultaneously transmit multiple images.

A. S. Ansarietal. [6] presentedan image Steganography algorithm that can work for cover images of multiple formats. Having a single algorithm for multiple image types provides several advantages. For example, we can apply uniform security policies across all image formats, we can adaptively select the most suitable cover image based on data length, network bandwidth and allowable distortions, etc. We present our algorithm based on the abstract concept of image components that can be adapted for JPEG, Bitmap, TIFF and PNG cover images. To the best of our knowledge, the proposed algorithm is the first Steganography algorithm that can work for multiple cover image formats. In addition, we have utilized concepts like capacity pre-estimation, adaptive partition schemes and data spreading to embed secret data with enhanced security. The proposed method is tested for robustness against Steganalysis with favorable results. Moreover, comparative results for the proposed algorithm are very promising for three different cover image formats.

P. Grandhe,etal. [7] aim is to generate a tool that will give out a benchmark value of how precisely the message is stored in the cover file. Using Stego and bulk analysis the information about the presence of the stego medium in the message can be known to the user. All these analysis methods make the tool more enhanced and secure.

In [8], batch steganography is used to secure data transmission from one end to the other. Often a password can be used for encoding the payload into the cover image. Here the data is encrypted using hashing and encryption techniques, SHA-256 and AES. The passwords used for encryption have been used after the logical operation XOR. Thus, the information has been encrypted twice using the XORed password for first and second input password for the next time. It increases the security of the data and makes the decryption almost impossible without knowing both the passwords and the encryption method. The encoded data is then embedded within the pixels of the original image using the LSB method. This prevents data theft and any possibilities of Man-in-the-Middle attacks since the time required for decrypting the data is drastically high without the knowledge of the inputs and the techniques used.

In [9], a more accurate image steganography method is proposed, where a multi-level feature fusion procedure based on GAN is designed. Firstly, convolution and pooling operations are added to the network for feature extraction. Then, short links are used to fuse multi-level feature information. Finally, the stego image is generated by confrontation learning between discriminator and generator.

Experimental results show that the proposed method has higher steganalysis security under the detection of high-dimensional feature steganalysis and neural network steganalysis. Comprehensive experiments show that the performance of the proposed method is better than ASDL-GAN and UT-GAN.

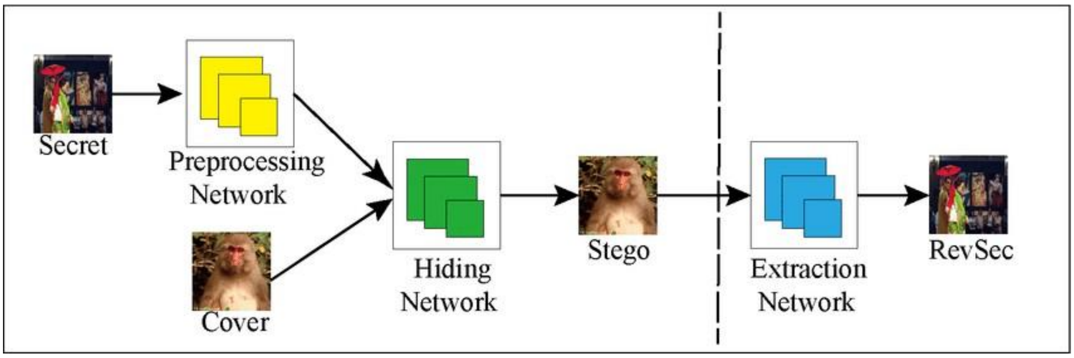
Zhao,etal[10]proposed a multi dilated generation countermeasure network (Multi Dilated GAN) model to improve the information steganography quality of images. Based on the discriminator of steganogan model, multiple convolution and expansion convolution are adopted. By selecting different expansion rates in the expansion convolution and matching with different receptive fields, different feature layers of the network can be effectively extracted; Multiple convolution processing is carried out on different feature layers to connect the features extracted from different directions, and distinguish the steganographic image from the carrier image, so that the discriminator can more accurately distinguish the difference between the steganographic image and the carrier image, so as to improve the steganographic ability of the encoder of the model. Experiments show that this model can effectively improve the steganogan model's steganographic accuracy and steganographic quality while maintaining a high steganographic capacity of 4.4bit/pixel.

In [12],authors proposed a new adversarial embedding scheme for image steganography. Unlike those related works, they first combine multiple gradients of cover and generated stegos to determine the directions of cost modifications. Next, instead of adjusting all or a random part of embedding costs in existing works, we carefully select the candidate costs according to the amplitudes of cover gradients and their costs. Extensive experimental results demonstrate that by adjusting a tiny part of embedding costs (less than 5% in most cases), the proposed method can significantly improve the security of five modern steganographic methods evaluated on both re-trained CNN-based and traditional steganalyzers, and achieve much better security performances compared with related methods. In addition, the security performances evaluated on different image database show that the generalization of the proposed method is good.

3. PROPOSED SYSTEM

Step 1: Steganography Dataset

The initial step in our research involves acquiring a suitable dataset for steganography. This dataset consists of various images that will serve as carriers for the hidden information. The selection of these images is crucial, as their characteristics can influence the effectiveness and imperceptibility of the steganographic technique. Typically, a diverse set of images with varying resolutions, color distributions, and complexities are chosen to ensure the robustness and generalizability of the proposed method.



www.ijbar.org
ISSN2249-3352(P)2278-0505(E)
CosmosImpactFactor-5.86

Fig. 1: Image steganography architecture.

Step 2: Dataset Preprocessing

Before utilizing the dataset for steganographic embedding, it is crucial to preprocess the images to ensure their suitability for the process. This preprocessing involves several steps to ensure the integrity and consistency of the data. First, null value removal is performed to eliminate any corrupt or incomplete images within the dataset. Any image files containing null values or missing data are either corrected or removed to maintain the quality of the dataset. Additionally, label encoding is applied to any categorical data associated with the images, such as image type or source. This step converts these categorical values into numerical representations, which is particularly useful when the dataset includes metadata that could impact the embedding process. By performing these preprocessing steps, we can ensure that the dataset is clean, consistent, and ready for secure data embedding.

Step 3: Existing LSB Substitution in Single Image Steganography Algorithm

To establish a baseline for performance comparison, we first implement the traditional Least Significant Bit (LSB) substitution technique. In LSB steganography, the least significant bit of each pixel in an image is replaced with the bits of the secret message. This method is straightforward and widely used due to its simplicity and ease of implementation. However, it is also more susceptible to detection and less robust against image manipulations and attacks. By implementing this method, we can evaluate its strengths and weaknesses and set a benchmark for our proposed approach.

Step 4: Proposed PVD in Multiple Image Steganography

The core of our research focuses on developing an advanced steganographic technique using Pixel Value Differencing (PVD) across multiple images. This technique follows a series of critical steps to ensure both security and imperceptibility. First, the secret message is divided into smaller chunks through message slicing, with each chunk embedded into a different image. This distribution of the message enhances security, as accessing the entire message requires obtaining all the carrier images. Next, the PVD algorithm is applied to each image, where the differences between adjacent pixel values are leveraged to embed bits of the secret message. This method is more effective and less noticeable than directly altering individual pixel values, making the embedded information harder to detect. Finally, the images with the embedded data are saved, ensuring that the information remains imperceptible to the naked eye and resilient against common image processing operations. This approach enhances the overall security and robustness of the steganographic technique.

Step 5: Performance Comparison

To evaluate the effectiveness of our proposed PVD-based multi-image steganography technique, we conduct a thorough performance comparison with the traditional LSB substitution method. This comparison focuses on several key aspects to ensure the superiority of our approach. First, we assess imperceptibility by evaluating the visual quality of the steganographic images to confirm that the embedded information remains unnoticed. Metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) are used to quantify this aspect. Second, we test the robustness of the embedded data, examining its resilience against various attacks and manipulations, including compression, resizing, and noise addition. This ensures that the embedded information remains intact even under adversarial conditions. Lastly, we compare the capacity of both techniques, measuring the amount of data that can be embedded without compromising the image quality. This comprehensive evaluation helps determine the overall effectiveness and advantages of the PVD-based method over traditional techniques like LSB substitution.

Step 6: Prediction of Output from Test Data

Finally, we develop and test a predictive model based on the PVD steganography technique. The process begins with training the model using a portion of the preprocessed dataset, allowing the model to learn how to embed and extract data effectively using the PVD technique. During this phase, the algorithm parameters are fine-tuned to optimize both the embedding and extraction processes. Once the model is trained, it is tested on a separate set of images to validate its performance. This testing phase focuses on evaluating the model's accuracy in correctly embedding and extracting the secret message, as well as its robustness against various image alterations. The results from the test data are then thoroughly analyzed to assess the model's overall effectiveness. This includes evaluating key factors such as imperceptibility, robustness, and capacity of the steganographic method, providing insights into how well the model performs in real-world scenarios.

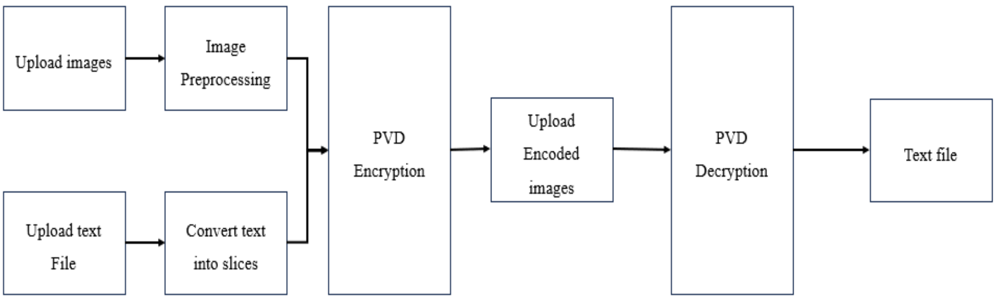


Figure 2: Architecture diagram of multiple image steganography

3.2 Pixel Value Differencing (PVD)

Pixel Value Differencing (PVD) is an innovative approach to steganography that operates in the spatial domain of images. While PVD is commonly applied to single images, its extension to multiple image steganography introduces new dimensions of security and capacity. The fundamental concept of PVD revolves around manipulating the pixel values of multiple images to embed hidden information.In PVD, the difference between the pixel values of adjacent pixels is utilized for data embedding. By carefully adjusting these differences, information can be hidden without significantly altering the visual appearance of the images. PVD operates in the spatial domain, making it resilient to frequency-based attacks. Unlike frequency domain techniques that might be susceptible to transforms, PVD directly modifies pixel values for data concealment.

Multiple Image Steganography

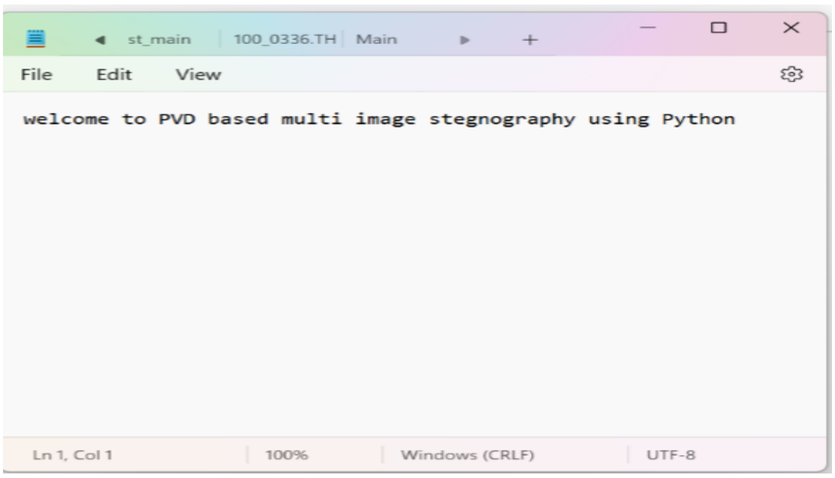
The extension of PVD to multiple images involves distributing hidden information across a set of images. This approach enhances security by dispersing the embedded data, making it more challenging for adversaries to detect or extract the complete message.

Embedding Process: The embedding process in PVD-based multiple image steganography follows a structured series of steps to ensure both security and effectiveness. First, a set of cover images is selected to act as carriers for the different segments of the hidden message. These images are carefully chosen to provide sufficient pixel data for embedding. Next, the hidden message is divided into segments, each corresponding to one of the selected cover images. Each segment is then embedded into its respective image. The PVD algorithm is applied to each image, where pixel values are adjusted based on the differences between adjacent pixels. These differences are manipulated to encode the hidden message in a way that is not easily detectable. To further enhance security, a secure key is integrated into the embedding process. This key governs how the pixel value differences are adjusted, making it extremely difficult to extract the hidden information without the key. This multi-layered approach ensures that the steganographic embedding is both secure and imperceptible.

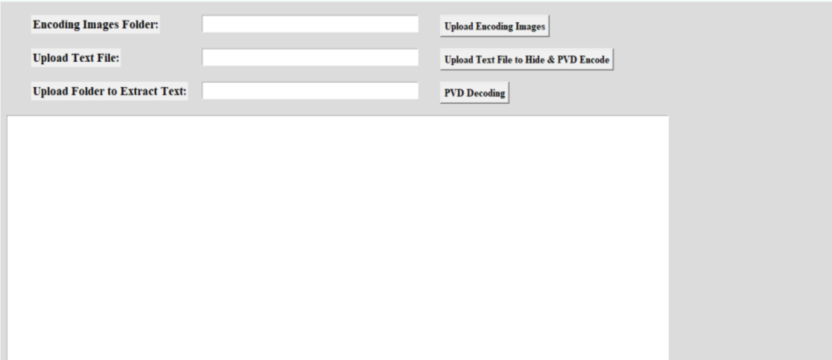
Extraction Process: The extraction process is designed to accurately retrieve the hidden message from the stego images through a series of systematic steps. First, the stego images, which contain the segments of the hidden message, are selected. The PVD algorithm is then applied in reverse to extract the differences between the pixel values of the images. These differences are carefully analyzed to reconstruct the segments of the hidden message. To ensure that the extraction is accurate and reliable, the secure key, which was used during the embedding process, is employed in the extraction phase. The key is crucial for correctly interpreting the pixel value differences and ensuring the precise retrieval of the hidden information. This method ensures that the embedded message is successfully recovered without compromising its security.

4. RESULTS AND DISCUSSION

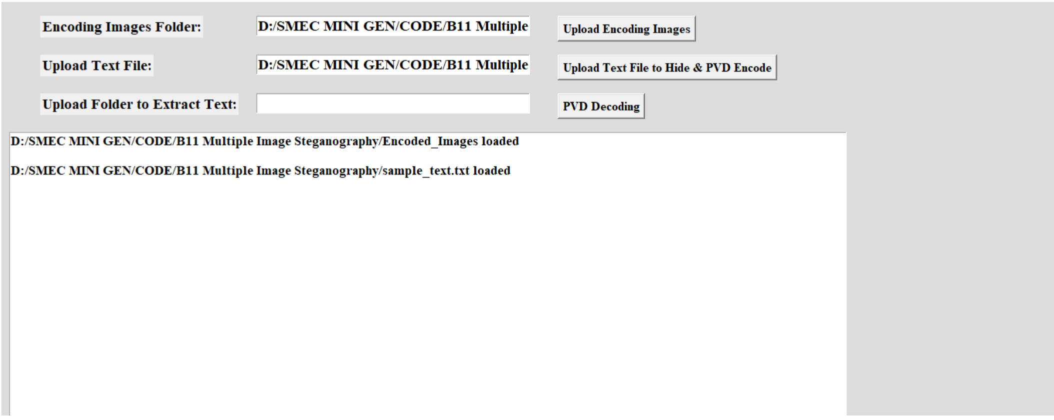
Improved File Security System Using Multiple Image Steganography. In this Research as per your instructions we have developed PVD (Pixel Value Differencing) based image steganography where user can upload multiple images folder and then upload text file which has to be slice and embed in all those uploaded images. All embed images will get saved inside 'Encoded_Images' folder with text slice data hidden inside it. While decoding we can upload desired folder from 'Encoded_Images' folder to extract text.To embed text we are using below sample text file



Above sample text will get sliced and hide inside multiple images



In above screen click on ‘Upload Encoding Images’ button to upload folder with multiple images like below screen



In above screen images uploaded and now click on ‘Upload Text File to Hide & PVD Encode’ button to upload sample text file to slice file and then embed in all images and get below output

In above screen selecting and uploading ‘images2’ message encoded folder and then click on ‘Select Folder’ button to extract message and get below output

5. CONCLUSION

Multiple Image Steganography represents a significant advancement in the field of covert communication and secure data transmission. The technique of distributing hidden information across a series of images, coupled with the Pixel Value Differencing (PVD) algorithm, offers a potent combination of security, resilience, and imperceptibility. The strategic division of data, error-correction techniques, spread spectrum methods, and secret sharing schemes contribute to the robustness and reliability of the steganographic system.The incorporation of cryptography and encryption enhances the confidentiality of the concealed information, while authentication and watermarking techniques provide mechanisms for verifying the integrity of the images. Hybrid approaches, integrating various steganographic methods and security measures, offer adaptability and versatility to meet diverse security requirements.However, the implementation of multiple image steganography is not without challenges. Synchronization during the embedding and extraction processes is crucial for accurate data retrieval. Striking a balance between usability and security is a delicate consideration, requiring thoughtful trade-offs to create an effective and user-friendly steganographic system.

REFERENCES

[1]. B. Sultan and M. A. Wani, "Multi-data Image Steganography using Generative Adversarial Networks," 2022 9th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2022, pp. 454-459, doi: 10.23919/INDIACom54597.2022.9763273**.**

[2]. X. Liao, J. Yin, M. Chen and Z. Qin, "Adaptive Payload Distribution in Multiple Images Steganography Based on Image Texture Features," in IEEE Transactions on Dependable and Secure Computing, vol. 19, no. 2, pp. 897-911, 1 March-April 2022, doi: 10.1109/TDSC.2020.3004708.

[3]. M. Srivastava, P. Dixit and S. Srivastava, "Data Hiding using Image Steganography," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-6, doi: 10.1109/ISCON57294.2023.10112069.

[4]. A. G. Benedict, "Improved File Security System Using Multiple Image Steganography," 2019 International Conference on Data Science and Communication (IconDSC), Bangalore, India, 2019, pp. 1-5, doi: 10.1109/IconDSC.2019.8816946.

[5]. S. Mukhopadhyay and H. Leung, "Multi Image Encryption and Steganography Based on Synchronization of Chaotic Lasers," 2013 IEEE International Conference on Systems, Man, and Cybernetics, Manchester, UK, 2013, pp. 4403-4408, doi: 10.1109/SMC.2013.751.

[6]. A. S. Ansari, M. S. Mohammadi and M. T. Parvez, "A Multiple-Format Steganography Algorithm for Color Images," in IEEE Access, vol. 8, pp. 83926-83939, 2020, doi: 10.1109/ACCESS.2020.2991130.

[7]. P. Grandhe, A. M. Reddy, K. Chillapalli, K. Koppera, M. Thambabathula and L. P. Reddy Surasani, "Improving The Hiding Capacity of Image Steganography with Stego-Analysis," 2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS), Raichur, India, 2023, pp. 01-06, doi: 10.1109/ICICACS57338.2023.10100146.

[8]. R. Joshi, A. K. Bairwa, V. Soni and S. Joshi, "Data Security Using Multiple Image Steganography and Hybrid Data Encryption Techniques," 2022 International Conference for Advancement in Technology (ICONAT), Goa, India, 2022, pp. 1-7, doi: 10.1109/ICONAT53423.2022.9725949.

[9]. Z. Wang, Z. Zhang and J. Jiang, "Multi-Feature Fusion based Image Steganography using GAN," 2021 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW), Wuhan, China, 2021, pp. 280-281, doi: 10.1109/ISSREW53611.2021.00079.

[10]. X. Zhao and H. Huang, "Research on Image Steganography Based on Multiple Expansion Generation Adversarial Network," 2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC), Greenville, SC, USA, 2021, pp. 361-366, doi: 10.1109/ICFTIC54370.2021.9647204.

[11]. B. Wei, X. Duan and H. Nam, "Image Steganography with Deep Learning Networks," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 1371-1374, doi: 10.1109/ICTC55196.2022.9952432.

[12]. M. Liu, W. Luo, P. Zheng and J. Huang, "A New Adversarial Embedding Method for Enhancing Image Steganography," in IEEE Transactions on Information Forensics and Security, vol. 16, pp. 4621-4634, 2021, doi: 10.1109/TIFS.2021.3111748.