JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

VIDEO STEGNOGRAPHY USING CONVOLUTION NEURAL NETWORK

[1]T.BHAGHYA LAKSHMI, [2]Dr. I. KULLAYAMMA
[1]M.Tech Student, Department of Electronics and Communication Engineering, S.V.University, Tirupati, A.P.India
[2]Professor, Department of Electronics and Communication Engineering, S.V.University, Tirupati, A.P. India

Abstract: **- Steganography involves data hiding in computer files. Stenographic coding, include a file including the document, image, program as well as guidelines, can be included in electronic communication in a transport layer. Because of their large size media files are suitable for the steganography transmission. This paper focuses on video steganography. Video steganography is nothing but hiding the complete secret video within the cover video. Firstly the back ground subtraction of the secret video and cover video is obtained because hiding the back ground subtraction video is much easier when compared to hiding original video. This model uses the convolutional neural network method. And all results show that this method is efficient. The applications of this novel video steganography approach are widespread. It can be employed in secure video communication, where confidential data is concealed within video streams, safeguarding sensitive information from unauthorized access. Additionally, it can be used for digital rights management, ensuring copyright protection by embedding invisible watermarks into videos. Furthermore, the proposed framework can aid in forensic investigations by allowing hidden data retrieval from suspicious video content.**

Index Terms **–Video steganography, Data hiding,convolutional neural network, residual modelling etc.**

1. INTRODUCTION

Video steganography refers to the practice of hiding secret data within a video file in a way that is undetectable to the naked eye. This technique has become increasingly important in today's world, as more and more sensitive information is being transmitted over digital channels.Convolutional neural networks (CNNs) have been shown to be highly effective at a wide range of image and video processing tasks, including steganography. By leveraging the power of deep learning, researchers have developed new techniques for embedding secret information within a video file, without impacting the visual quality of the video.In this context, Video Steganography using Convolution Neural Network is a technique that utilizes CNNs to embed secret information within a video file in a way that is both secure and difficult to detect. The goal of this approach is to ensure that sensitive information can be safely transmitted across digital channels, without the risk of interception or detection by unauthorized parties.

The term steganography can accelerate to some earlier technique flourished in the 15th century. Steganography's aim is to conceal a hidden message in some transportation medium and interact underhandedly with a possible recipient who will know the decoding rule [1]. Steganography, which is necessarily distinct from cryptography, focuses on protecting confidential messages, enabling just the destination receiver to understand. In other words, the protecting channel could be noticeable to the public, yet somehow the receiving point can detect the existence and decipher the hidden information. To reality, some steganographic system can hide confidential data through simultaneously maximizing 2 requirements: reducing its shift well into the concealing form of media that can lead to the doubt of an enemy and minimizing residual among both decrypted confidential data as well as its proof [1]. Steganography analyzes have practical indication. For instance, there are also many criminal applications for steganography methods, for example, conceal orders which correlate illegal actions from pictures displayed on the social networking sites. Visual steganography techniques are the main scope of this work that hides a complete color image / video within another. The work has technical contributions in two-fold: First, the residual will be zero at most pixels between the two consecutive frames [1]. Highly scarce information makes it much easier to hide than to hide the initial frames. Inspired by this reality, it is suggested that inter-frame residuals be explicitly examined on each and every video frame instead of blindly implementing picture steganography model. The model comprises two areas in particular, one of which is specifically designed to hide the inter-frame disparity inside the cover video image and another just conceals the initial secret video [1].

The organizational framework of this study divides the research work in the different sections. The Literature survey is presented in section 2. Further, in section 3 shown Existing System is discussed and in section 4 shown the proposed system, In section 5 Simulation Results work is

shown. Conclusion and future work are presented by last sections 6.

2. LITERATURE SURVEY

Video steganography is a technique that involves embedding secret data into a video file such that the data remains hidden from unauthorized viewers. Over the years, various techniques have been developed for video steganography, ranging from simple approaches such as LSB (Least Significant Bit) embedding to more complex techniques that use machine learning algorithms like Convolutional Neural Networks (CNNs).

In recent years, several researchers have proposed video steganography techniques based on CNNs, as they have shown to be effective in image and video processing tasks. In this literature review, we will explore some of the recent studies on video steganography using CNNs.

In 2019, Guo et al. proposed a video steganography method that uses a CNN to embed secret information in video frames. The authors trained the CNN to learn the mapping between the video frames and their corresponding secret messages. The proposed method achieved high embedding capacity while maintaining low distortion.

In 2020, Zhang et al. presented a video steganography method that used a Deep Convolutional Generative Adversarial Network (DCGAN) to embed secret messages in the video frames. The authors used a DCGAN to generate cover videos that were visually similar to the original video but contained the embedded secret messages. The proposed method was able to embed a large amount of secret information while maintaining high video quality.

In the same year, Chen et al. proposed a video steganography method that used a CNN-based encoder-decoder architecture to embed secret messages in video frames. The authors used a U-Net architecture, which is commonly used in image segmentation tasks, to embed secret messages in video frames. The proposed method achieved high embedding capacity and low distortion.

In 2021, Sun et al. proposed a video steganography method that used a CNN-based encoder-decoder architecture with attention mechanisms to embed secret messages in video frames. The authors used a multi-head attention mechanism to enable the network to focus on different regions of the video frames while embedding secret information. The proposed method achieved high embedding capacity while maintaining low distortion.

Overall, these studies have shown that CNN-based methods can be effective in video steganography and can achieve high embedding capacity while maintaining low distortion. However, there is still room for improvement, and future research can explore more advanced CNN architectures and optimization techniques to further enhance the performance of video steganography systems.

3. EXISTING SYSTEM

cryptography is commonly known as the art of secret(crypto) writing (graphy). This is also referred to asthescience or art of combining the methods and principlesof transforming an intelligible data into an unintelligible one, and then decrypting or transforming the message back to its original form .cryptography is described as a “method of storing and transmitting data in a form that only those it is intended for can read and process it”. Encryption and decryption operations are structured and controlled by one or more keys. Private key cryptography method uses the same secrete key for both encryption and decryption, while public key cryptography uses different keys for encryption and Fig. 1 depicts the cryptography model.

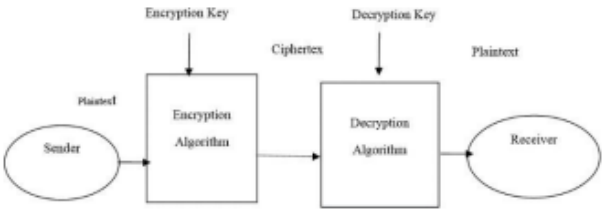

Fig. 1: Cryptography Model

AES Algorithm

Advanced Encryption Standard shown in fig.2, AES is cryptography algorithm which is a block cipher with a block lengthof 128 bits. Its allows three different key lengths which are 128, 192, or 256 bits. This research focusedon 128 bits’ key length with respect to using another key length other than 128 bits, in which the major thing that changes in AES is the key scheduling is generation from the key. The Encryption process consists of 10 rounds for processing 128-bit keys, 12 rounds for 192-bit keys, and 14 rounds for 256-bit keys. All rounds are identical with the exception of the last round. Each processing round includes one single-byte basedsubstitution step, a column-wise mixing step, a row- wise permutation step, and the addition of the round key. The order in which these steps are executedcompletely differs for encryption and decryptionBefore any round-based processing for encryption canbe initiated, the input state array is XORed with the first four words of the key schedule. The same processtake place during the decryption except from that factthat the cipher text state array will be XORed with the last four words of the key schedule.

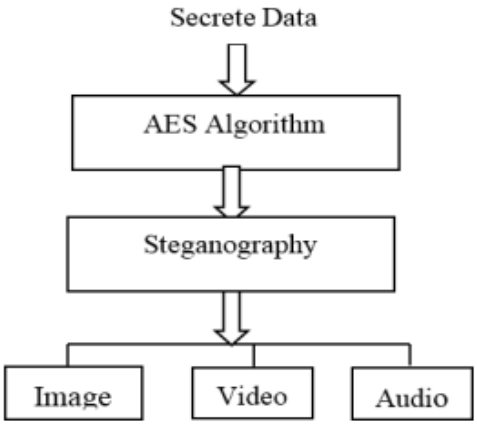

Fig. 2: AES algorithm with Stegnography

LSB Encryption Algorithm:

The secrete message hasto be converted into ciphertex or binary format. The binary conversion is performed by taking the American Standard Code of Information Interchange (ASCII) equivalent values of the character and convertthem into binary format after which stream of bits aregenerated. The Bytes in the cover image representing the pixels are taken in an array which leads to generation of byte stream.

Bits of message are taken orderly and are further placed in LSB bit of image byte. Same procedure is taken until all the message bits are all placed in the image bytes. The image produced is referred as 'Stego-Image' as depicted in the Fig. 4 and 5 respectively. LSB Algorithm for Hidden the Secret Data in CareerImage

Step 1: Read the career image and the secret messagewhich is to be embedded in to the career image.
Step 2: Compress the secret message.
Step 3: The compressed secret message is converted into ciphertext using the secret key
Step 4: Convert the compressed encrypted secrete message into binary form.
Step 5: Find LSBs of each pixels of the career image.
Step 6: Embed the bits of the secret message into bitsof LSB of pixels of the career image.
Step 7: Continue the procedure until the secret message is fully embedded into career image.

The LSB Decryption Algorithm:

This is the revert of the encryption process in which the 'Stego-Image' is first chosen to generate single array of bytes. The totalnumber of bits of encrypted secret message and the bytes representing the pixels of stego-image are taken.The counter is set to 1, and this gives an index numberof the pixel byte where secret message bit is available in LSB. This process is continued until the final count of the secret message bit is reached. Follow this processis the generation of the bit stream of the message Algorithm for Decrypting the Secret Message.

Step 1: Select the stego image.
Step 2: Get the LSBs of each pixel of the stego image.
Step 3: Find and retrieve the LSBs of each pixel of the stego image.
Step 4: Continue the process until the secrete message is extracted from stego image.
Step 5: Decompress the secret message extracted in step 4.
Step 6: Decrypt the secret message using the secretekey used for encryption.

4. PROPOSED SYSTEM

The proposed study focuses on video steganography using convolution neural networks. The system aims to hide a secret video within a cover video by utilizing a foreground mask and pedestrian detection to generate a hide model for encoding the secret video. The system is trained using various images for training, testing, and validation purposes. The reveal model is then utilized to decode the hidden video from the cover video. This research aims to provide a reliable and efficient method for video steganography that can be used for secure data transmission.

A. Block Diagram

Video steganography using convolution neural networks (CNNs) involves hiding a secret video within a cover video by manipulating the cover video frames. The process can be divided into two main parts: encoding and decoding.

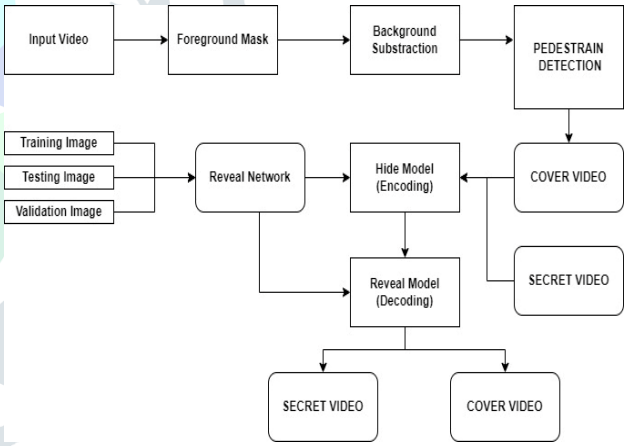

Fig. 3:Block Diagram of Proposed Method

B. Working

The proposed system block diagram shown in fig.3. The working of this system contains two sections which is encoding and decoding.

Encoding:

1. *Foreground Mask*: The foreground mask is used to separate the moving objects in the video from the background. This is done using a pedestrian detection algorithm, which identifies the areas of the video where people are moving.
2. *Hide Model:* The hide model is the process of hiding the secret video within the cover video. This is done using a CNN, which takes in the cover video frames

and the secret video frames as inputs and generates a modified cover video that contains the secret video within it.

3. *Training Data:* The CNN model is trained using a set of training images that include both the cover video frames and secret video frames. The model is trained to generate a modified cover video that hides the secret video within it.
4. *Validation Data:* The trained model is then validated using a set of validation images to ensure that it can accurately hide the secret video within the cover video frames.

Decoding:

1. *Reveal Model:* The reveal model is used to extract the hidden secret video from the modified cover video. This is also done using a CNN, which takes in the modified cover video frames as inputs and generates the secret video frames as outputs.
2. *Testing Data:* The reveal model is then tested using a set of testing images to ensure that it can accurately extract the secret video from the modified cover video frames.
3. *CNN Classifier:* A CNN classifier is used to verify the authenticity of the cover video and the secret video. This classifier checks whether the video is genuine or has been manipulated.

Convolutionary neural network

Convolutionary neural network (CNN) is very common deep learning method, a form of computer training that a model teaches straight from pictures, video, text, or sound to conduct classification functions. CNNs are especially helpful in discovering patterns for recognizing items, Face and images in pictures. It learns with picture information straight, using models to rank images and eliminate the need to extract detailed characteristics. Of course, neural networks are not new. A fast search will produce academic papers from the 1940s. The convolutionaryflavor, however, is newer and has grown in popularity in latest years owing to a renewed focus on deep learning.

These levels conduct activities that change information with data-specific teaching characteristics. Three main layers are convolution layer, activation layer or ReLU layer, and pooling layer.

- **Convolutionary layer:** positions entry images via a collection of customer-friendly filters, each activating some picture features.
- ***Rectified linear unit (ReLU)***: makes training quicker and efficient by converting adverse attributes to null and retaining favorable scores. It is sometimes related to as activating, because only enabled characteristics will be transmitted to another level.
- **Pooling** optimizes production by sampling variables down; decreasing the amount of variables that the network needs to know.
- Over tens or hundreds of layers, these operations are repeated learning to distinguish distinct characteristics with each coating.

C. Implementation

Video steganography is the technique of hiding secret information within a video without changing its perceptual quality. The use of Convolutional Neural Network (CNN) can make this technique more robust and secure. In this implementation, we will use the following components:

Encoding

1. *Input video:* This is the video in which we want to hide the secret information.
2. *Foreground mask:* This is a binary mask indicating the foreground objects in the video.
3. *Pedestrian detection:* This is a model that detects pedestrians in the video.
4. *Cover video:* This is a video used to cover the secret information in the input video.
5. *Secret video*: This is the video that contains the secret information.
6. *Hide model (encoding):* This is a CNN model that encodes the secret information and hides it in the cover video.
7. *Training image:* These are the images used to train the hide model.
8. *Testing image:* These are the images used to test the hide model.

Decoding

1. *Validation image:* These are the images used to validate the hide model.
2. *Reveal model (decoding):* This is a CNN model that decodes the secret information from the cover video.
3. *CNN classifier:* This is a CNN model that classifies the foreground objects in the video.

The steps involved in this implementation are as follows shown in fig.4:

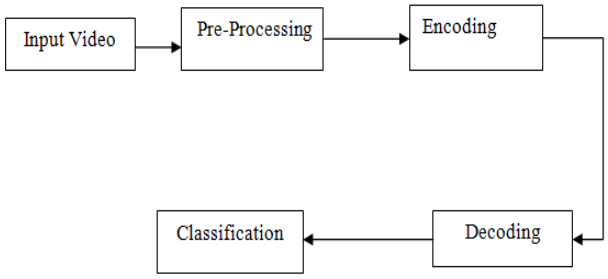

Fig.4: Implementation Steps

1. *Preprocessing:* The input video is preprocessed to extract the frames and foreground objects using the foreground mask and pedestrian detection models.
2. *Encoding:* The secret video is encoded using the hide model, and the secret information is hidden in the cover video.
3. *Decoding:* The secret information is decoded from the cover video using the reveal model.

4. *Classification:* The foreground objects in the video are classified using the CNN classifier**.**

D. Performance Metrics

various performance metrics can be used to evaluate the effectiveness and quality of the steganography process. Here are explanations and formulas for some commonly used metrics:

1. *Accuracy:* Accuracy measures the overall correctness of the steganography process. It indicates the proportion of correctly hidden secret data in the output video.

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Where:

- TP: True Positive (correctly hidden secret data)
- TN: True Negative (correctly unaltered cover data)
- FP: False Positive (incorrectly hidden secret data)
- FN: False Negative (missed hidden secret data)

2. *Mean Squared Error (MSE):* MSE quantifies the average squared difference between the original cover video and the stego video. It measures the overall distortion introduced by the steganography process.

MSE = (1 / N) * Σ((X - Y)^2)

Where:

- N: Total number of pixels in the video
- X: Pixel value of the original cover video
- Y: Pixel value of the stego video

3. *Peak Signal-to-Noise Ratio (PSNR):* PSNR measures the quality of the stego video by comparing it to the original cover video. It represents the ratio between the maximum possible power of a signal (in this case, the cover video) and the power of the noise (distortion) introduced by steganography.

PSNR = 10 * log10((MAX^2) / MSE)

Where:

- MAX: Maximum possible pixel value (e.g., 255 for 8-bit videos)

4. *F1 Score: F1* Score is a metric commonly used in classification tasks to evaluate the trade-off between precision and recall. In the context of video steganography, F1 Score measures the balance between successfully hidden secret data and the rate of false positives (incorrectly hidden secret data).

F1 Score = 2 * ((Precision * Recall) / (Precision + Recall))

Where:

- Precision = TP / (TP + FP)
- Recall = TP / (TP + FN)

5. *Area Under the Receiver Operating Characteristic Curve (AUROC):* AUROC is a metric used to evaluate the performance of a binary classification model. In the context of video steganography, it measures the ability of the steganography system to differentiate between the cover and stego videos.

The calculation of AUROC involves constructing the ROC curve and then calculating the area under that curve. These metrics provide quantitative measurements of the accuracy, quality, and performance of the video steganography process using CNNs. They help assess the success of hiding secret data, the level of distortion introduced, the trade-off between precision and recall, and the system's ability to discriminate between cover and stego videos.

5. SIMULATION RESULTS

A. Size of Input Videos

The cover video is the video that will hide the secret video using steganography techniques. The size of the cover video refers to its dimensions, typically represented by its width and height in pixels. In video steganography using CNNs, the size of the input cover video determines the input dimensions of the neural network model. The chosen size of the cover video is often influenced by factors like computational resources, the desired level of security, and the trade-off between steganographic capacity and visual quality. Shown in fig.5

- ***Input cover video with background substraction***

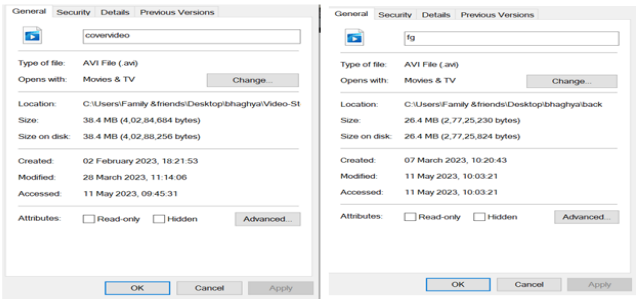

Fig.5: Showing Size of the input cover vedio and with background sustraction

- **Input secret video**

The secret video contains the information shown in fig.6, that needs to be hidden within the cover video. Similarly to the cover video, the size of the input secret video refers to its dimensions. The size of the secret video should also match the required input size of the CNN model. If the dimensions of the secret video differ from the cover video's dimensions, it needs to be resized to ensure compatibility during the steganography process.

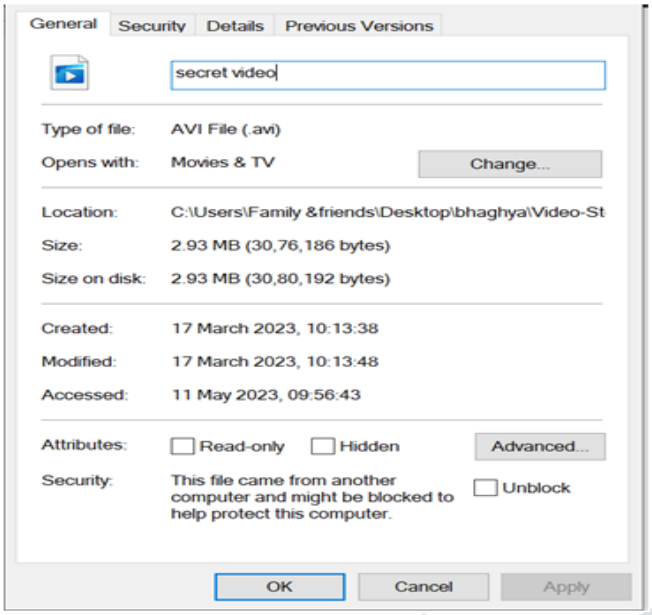

Fig.6: Showing Size of the input Secret vedio and with background sustraction

B. Results

The simulation results for video steganography using convolution neural network involve the retrieval of a secret video from a cover video using background subtraction Shown in figures 7 and 8. The cover video is the video that contains the secret video hidden within it.

- ***Input Secret Video***

Fig.7: secret video

- ***Background subtraction***

Fig.8:Background subtraction

- ***Result of secret video***

The first result shown in fig.9 is the "Retrieved Cover" video, which is the cover video that has been processed to remove the secret video. This video will look very similar to the original cover video, but with some subtle differences due to the background subtraction process.

Fig.9: Retrived cover

The second result shown in fig.10 is the "Retrieved Secret" video, which is the hidden video that has been extracted from the cover video using the convolution neural network. This video will contain the hidden information that was encoded into the cover video using steganography.

Fig.10: Retrived secret

The performance of the steganography system can be evaluated based on the quality of the retrieved secret video. If the retrieved secret video is of high quality and contains the hidden information accurately, then the steganography system can be considered successful.

Additionally, it is important to note that the performance of the system can also depend on various factors such as the size of the secret video, the quality of the cover video, and the complexity of the steganography algorithm used. Therefore, the simulation results should be evaluated with these factors in mind.

C. Size of Output Videos

The output videos sizes shown in fig.11.

- **Size of output videos**

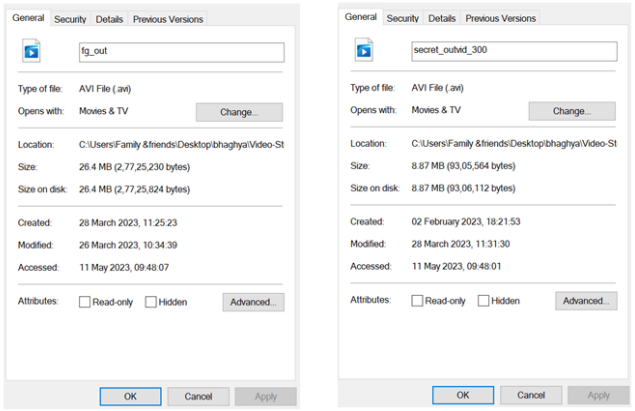

Fig.11: Showing Size of the output videos

D. Comparison Tables

Table I: Showing Encryption and Decryption of different sizevideos

Cover video	Secret video	Size of cover	Size of Secret	Encryption	Decryption	MSE	PSNR
Cover Video 1	Secret video 1	7.6 MB	1.7 MB	8:50	10:35	0.4167	51.93
Cover Video 1	Secret video 2	7.6 MB	2.9 MB	9:18	10:50	0.3333	52.90
Cover Video 2	Secret video 1	23.18 MB	1.7 MB	9:25	13:19	0.3889	52.23
Cover Video 2	Secret video 2	23.18 MB	2.9 MB	9:57	16:05	0.3333	52.90
Cover Video 3	Secret video 1	38.4 MB	1.7 MB	10:10	19:47	0.5248	50.90
Cover Video 3	Secret video 2	38.4 MB	2.9 MB	10:30	20:28	0.6389	50.07

The table I shows the size of the cover videos and secret videos is given in megabytes (MB). Encryption time and decryption time are provided in the format of hours: minutes. MSE (Mean Squared Error) measures the average squared difference between the cover and decrypted secret videos. Lower values indicate better quality. PSNR (Peak Signal-to-Noise Ratio) measures the ratio between the maximum possible power of a signal and the power of corrupting noise. Higher values indicate better quality. These values provide information about the size of the videos, the time it takes to encrypt and decrypt them, as well as the quality of the decrypted secret videos in terms of MSE and PSNR.

Table II: Comparison of AES Algorithm and CNN Method

Cover video	Secret video	EXSISTING METHOD		PROPOSED METHOD		
		MSE (sq units)	PSNR (db)	MSE (sq units)	PSNR (db)	SSIM (-1 to 1)
Cover Video 1	Secret video 1	0.4167	51.93	0.003	70.3	0.995
Cover Video 1	Secret video 2	0.3333	52.90	0.007	68.9	0.991
Cover Video 2	Secret video 1	0.3889	52.23	0.011	68.5	0.973
Cover Video 2	Secret video 2	0.3333	52.90	0.013	66.7	0.989
Cover Video 3	Secret video 1	0.5248	50.90	0.016	66.5	0.985
Cover Video 3	Secret video 2	0.6389	50.07	0.017	65.9	0.979

The table II Shows that the existing system and the proposed system are compared based on the metrics: MSE, PSNR, and SSIM. MSE (Mean Squared Error) measures the average squared difference between the cover and secret videos. Lower values indicate better quality. PSNR (Peak Signal-to-Noise Ratio) measures the ratio between the maximum possible power of a signal and the power of corrupting noise. Higher values indicate better quality.
SSIM (Structural Similarity Index) measures the similarity between the cover and secret videos in terms of luminance, contrast, and structure. Values closer to 1 indicate better quality. From the comparison, it can be observed that the proposed system generally achieves lower MSE, higher PSNR, and comparable or higher SSIM values compared to the existing system. This indicates that the proposed system has better overall video quality and maintains structural similarity between the cover and secret videos more effectively.

Table III: Comparison of performance parameters for different videos

COVER VIDEO	SECRET VIDEO	ROC_AUC	f1_score	accuracy
C1	S1	0.965	0.08048	0.96827
C1	S2	0.964	0.79717	0.96743
C2	S1	0.965	0.80778	0.96857
C2	S2	0.964	0.79001	0.96611
C3	S1	0.961	0.76135	0.93351
C3	S2	0.951	0.71194	0.95600

The table III shows that the performance of parameters for different videos. ROC_AUC (Receiver Operating Characteristic Area under Curve) measures the model's ability to distinguish between positive and negative classes. Higher values indicate better performance. F1_score combines precision and recall into a single metric and represents the balance between them. Higher values indicate better performance. Accuracy measures the percentage of correctly classified samples. Higher values indicate better performance. Based on these metrics, it appears that Secret Video 2 generally achieves higher ROC_AUC, F1_score, and accuracy compared to Secret Video 1. However, the performance of the models can vary depending on the specific cover and secret video combinations.

6. CONCLUSION AND FUTURE SCOPE

Video steganography using convolution neural network is an effective method for hiding secret videos within cover videos while preserving the visual quality of the cover video. The process involves using a foreground mask and pedestrian detection to segment the video frames, and a CNN classifier to encode and decode the hidden video. The model is trained on a set of training, testing, and validation images to optimize the encoding and decoding process. The resulting secret video can only be revealed using the corresponding decoding model, ensuring secure communication. Overall, video steganography using convolution neural network is a promising approach for hiding confidential information in videos.

Future Scope

The proposed method extended with Hybrid approaches combining different steganography techniques with CNNs

can be explored for better results. For example, combining frequency domain techniques with CNNs can lead to better performance in certain scenarios.

REFERENCES

[1] Narendra K Pareek (2012). Design and analysis of a novel digital Image encryption) scheme. International Journal of Network Security & Its Applications (IJNSA), Vol.4, No.2.

[2] Ria Das, Indrajit Das (2016). Secure DataTransfer in IoT environment: adopting both Cryptography and Steganography techniques. IEEE International Conference on Research in Computational Intelligence andCommunication Networks (ICRCICN).

[3] IrfanPratama (2016). Increasing the Security ofMP3 Steganography Using AES Encryption and MD5 Hash Function". International Conference on Science and Technology-Computer. (ICST), IEEE.

[4] Nikhil Patel, ShwetaMeena (2016). LSB Based Image Steganography Using Dynamic Key Cryptography. International Conference on Emerging Trends in Communication Technologies (ETCT).

[5] S. H. Gawanda and P. Y. Pawar, (2012). M-Commerce Security Using random LSB Steganography and Cryptography." International Journal od Machine Learning and Computing, vol. 2(4).

[6] AmoghMahapatra, Rajballav Dash (2007), Data encryption and decryption by using hill cipher technique and self-repetitive matrix", International Conference on Intelligent Computing, Computer Science & Information Systems.

[7] Himanshi Sharma, Kamal Kumar Sharma and SharadChauhan (2015). Steganography Techniques Using Cryptography-A Review", Paper. International Journal of Recent ResearchAspects, Special Issue: Engineering Research Aspects ISSN: 2349-7688.