Enhancing Data Security in video Steganography using Face Recognition

Alaknanda S. Patil[1†] and Dr. G. Sundari[2††]

Department of ECE, Sathyabama Institute of science and technology, Chennai, India

Summary
In today's modern digitized world, the proliferation of communication technology and the internet has considerably increased the transmission of data in large amounts. Hence, secure data transmission is becoming a pre-requisite. Digital Steganography provides the capability to protect private communication in today's Internet era. Steganography is a practice to shelter and conceal multimedia information in a disguised manner, or we can say it is the study of invisible communication. Steganography uses image, text, video, and audio to disguise secret information. The presented system is an amalgamation of face recognition and LSB based steganography with spread spectrum technology. In the encryption stage, the extracting the audio from the original multimedia video file. Then secrete message is inserted and embedded in the extracted audio. The audio steganography uses the Least Significant Bit (LSB) substitution algorithm with a spread spectrum technique to add the secrete audio. The processed audio is called stego audio. The stego audio is embedded back into the video frames. Along with this, attach a face image of the authorized user is to the original video. The processed video is transmitted over the communication channel. In the decryption stage, the face image is extracted from the image and recognized by using the CNN algorithm. The audio can then be decrypted using a 4-bit LSB decryption algorithm only if the two previous facial images match. Thus the introduction of face recognition makes the system more robust. Testing of the sustainability of the system, the statistical analysis of the embedding algorithm has been performed and checked that it not only maintains the visual similarity in the stego file but also leaves the other statistics unchanged after embedding. The performance of the proposed algorithm is assessed by PSNR, SSIM, and RMSE metrics.

Key words:
Audio; Face recognition; Information Security; LSB; Steganography; Video.

1. Introduction

These days' communication of data is carried out in all forms of life. Therefore, maintaining the safety and secrecy of data comes into play. It needs to form policies to avert, discern, document, and stand dangers to the information present on the internet. Information security policies and processes commonly involve both physical and digital protection for the data from unauthorized use, access, reproduction, or demolition [1]. These measures can include network intrusion detection systems, encryption key management, and many other techniques. The different commonly used method for secure communication is steganography, which is the knack of hiding messages so that the secret information is not visible to the third party and not identified quickly. It conceals valuable information in the general data, and the resultant (stego object) must resemble the original form. The prime motive of steganography is to resist any suspicion by the third party, and if any doubt is elevated, then the algorithm is of no use. Steganography takes into account the human visual system (HVS), which cannot recognize slight distortion in the cover object. Thus steganography is a method of sharing secret information by making it inconspicuous to non-authenticated users. The term 'steganography' exemplifies the act of subtly inserting data into digital mediums together with video, image, and sound records.

Steganography (Greek word stegano meaning "covered" and graphic or "writing") is brought to use to hide messages into a more sophisticated type of information such as images, audio, or videos. These are called mediums. Steganography takes advantage of the information that humans cannot notice the minor changes in the medium. Following are some of the terms related to steganography:

- Secret data: The data that needs to send covertly from one place to another.
- Cover Medium: It refers to the medium that is used to cover the secret data.
- Stego Object: It refers to the cover medium once the secret has been successfully embedded.
- Stego Key: This key defines how the data is inserted or embedded within the cover medium. It is used at both the embedding and retrieving process.
- Imperceptibility: It defines the quality of the stego object. The imperceptibility is nothing but the undetectable nature of the cover medium once the secret data is embedded.
- Capacity: The amount of data that is possible to be inserted or embedded into the cover medium and stego object remains undetectable at the same time.

The main aim of the steganography is to deliver the information very secretly by using other mediums like image, audio, or video. These mediums are called cover objects of the steganography. The secret message can be image, text,

Manuscript received August 5, 2020
Manuscript revised August 20, 2020

audio, or video, called a message object. The embedding message object and cover object forms the stego-object.
There are five types of steganography according to the cover object used is as shown in Fig.1.
Text Steganography deals mainly with concealing text in Text Files, and Binary Files [2]. Text within a video may be obscured or hidden in any way. Text steganography is very capable of concealing text data in Cover Text Data.

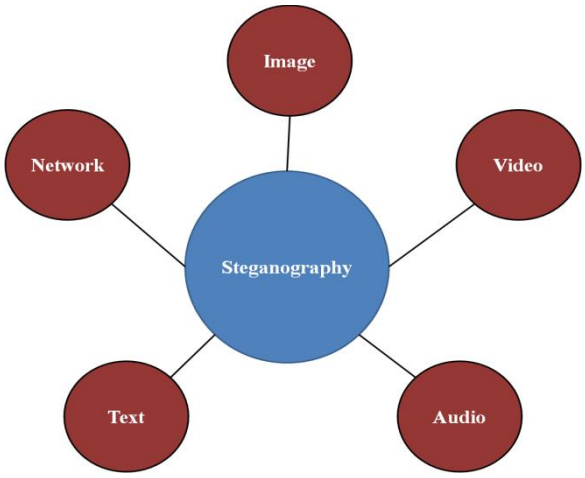


Fig. 1 Types of Steganography

The steganography of digital images mainly deals with the concealment of data inside a cover image [3]. They are very common in the Internet age nowadays. In steganography, digital photographs are considered highly used to cover media [4]. You may label Digital Image a set of pixels. Pixels are selected to hide data, based on their intensities.
Video can be defined as the integration of audio and still image collection, which moves in a sequence of the time constant. Due to high embedding payload than a digital image [5], videos are becoming popular as a cover object in steganography, and temporal features of a video also provide perpetual redundancy, which is not available in digital images. Because of the availability of a large number of frames, hidden data inside a video could easily be concealed.
A steganographic system is comprised of two algorithms, the first is for hiding, and the second is for retrieving. The hiding process is about embedding data within the cover medium. Therefore, this process should be constructed carefully to be sure the stego object is identical to the cover medium as possible, which makes sure that the presence of the message is undetectable. Therefore, the components of the embedding process system consist of a secret message and a cover medium as inputs, a steganography algorithm as the method of hiding, and a resulting stego object as the output. A secret key can be used for hiding the data as a third input. It will increase the robustness, and safety of the hidden data, such that there is no way the data is retrieved in the nonappearance of the secret key even though the algorithm of hiding is known.
Thus using steganography, the secrete message can be embedded inside the other information and sent to anyone without knowing the existence of the secret message. Hiding data in the audio is less doubtful than communicating an encrypted file.
In this proposed approach, initially, the audio from the video (.avi, .mpeg, or .ogg format) file is extracted called original audio [5]. The secrete audio (of format .wav, .mp3, .aiff, etc.) is embedded into the original audio with the aid of the Least Significant Bit (LSB) method with spread spectrum technology. The stego audio is again embedded with the video file. This embedded video is transmitted over the secure communication channel. At the receiver side, stego audio is extracted from the embedded video file, and the secrete audio from the stego audio is decrypted using the LSB method. This method provides two-stage security.
Face recognition is used along with LSB by building a more robust system. Face recognition is done before the audio-video decryption process to ensure that the received message is sent from the desired user.
The proposed paper is organized as follows; Section 2 offers an overview of the recent development in audio-video steganography using different algorithms and their advantages and disadvantages. Section 3 presents the proposed methodology for the two-stage steganography approach. Section 4 demonstrates the results qualitatively and quantitatively. Lastly, the conclusion is given in Section 5.

2. Literature Survey

There are various approaches implemented for steganography. Development can be done after reviewing existing progress in the steganography.
Moresh et al. [7] presented a mixture of image hiding and image encryption methodology to offer high security to the data/message for communication over the channel. The essential purpose of this technique is to deliver high-level security. Blowfish algorithm and LSB technique are used for encryption. The system performance is examined with PSNR and MSE.
Save et al. [8] proposed the use of audio-video stenography for data security. The approach was focused on hiding the secret information in arrears of audio and the recipient's face image of the video. RSA and LSB Algorithm are employed for hiding the secret image and text, and face recognition is accomplished using the PCA algorithm. Embedding image and text into the video and audio file and merging them into a stego file on the transmitter side and then face recognition technology is applied at the receiver side to ensure the security by approving the recipient.

S. M. Masud Karim et al. [9] familiarized with LSB based steganography system wherein the secret message is hidden within the cover image in non-sequential order with an encryption key for improved security. The secret keyword is transformed into ASCII stream and then into a binary stream. The stream of binary bits then XORed with the binary bits of an image's red color plane. The approach achieved a high PSNR value proving it to be more accurate.

Nadeem et al. [10] demonstrated the steganography approach with the use of LSB. This technique is additionally enhanced by a bit inversion technology to improve the quality of the stego-image. In this method, in place of traditional sequential embedding, randomly hiding the data bits in LSB has been presented. The LSBs of the stego image are inverted after embedding to lessen the number of reformed LSBs. This procedure scatters the bit in the cover-image and thus makes it difficult for an invader to regain the original message. This approach shows good PSNR along with image quality.

Murugan et al. [11] offered the steganography approach for jpeg and AVI format video by exploiting the swap technique. This technique delivers a simple algorithm with intricate encryption. They recommended that the use of UTF-32 encoding along with swapping algorithms can enhance the métier of steganography with lesser distortion and capability. The experimentation results demonstrated that the output of AVI steganography was improved without loss of size, data, and quality of the original video.

S. Kumar Et al. [12] proposed a method of steganography focused on the novel detection of fuzzy edges, which inserts secret information in gray images without significant modification in the cover images. The technique presented effectively detects the sharp contours of cover images that are used to embed secret bits of information. After inserting the code, the edges of the image will be retained to precisely retrieve the data at the intended recipient.

Mukesh Dalal et al. [13] proposed an approach for embedding secret data in a video sequence using the LSB technique with object detection. The embedding venue was selected by utilizing background subtraction and blob analysis, object detection method using GMM. The proposed method provides more security and imperceptibility as the data was embedded in the moving objects, and the changes in the moving objects are difficult to notice rather than the static region in a video.

Kamran et al. [14] presented the data security method with the LSB algorithm, also referred to as a Distributed LSB algorithm. In this method, the data is hidden plane by plane in a random manner. The amount of data to be hidden is decided to depending upon the grayscale intensity value of the pixel. It makes the algorithm vigorous and more operative in terms of data hiding volume and degradation of the cover image.

Yugeshwari Kakde et al. [15] developed a system of audio-video steganography by grouping the steganography of audio and image. This technique uses a Singular Value Decomposition (SVD) and Discrete Wavelet Transform (DWT) to cover text in the audio file. It used the method of random LSB audio steganography.

Maleki et al. [16] suggested methods of hiding adaptive and non-adaptive data for grayscale images that depend on the function of the modulus. The adaptive scheme addresses the notion of visual awareness in humans and a non-adaptive scheme which, in terms of security level and embedding capability, generates the least distortion for the stego-image. There is greater visual consistency to the non-adaptive algorithm.

Sumanth et al. [17] proposed the image and audio steganography with face recognition methodology for authentication. LSB is used for the image as well as audio steganography, and face recognition is accomplished using the PCA algorithm. Text information is hidden into audio archives productively; furthermore, deduce the audio file and focused on extracting secret text.

Kaushik et al. [18] proposed a system that is a combination of face recognition and steganography. The user has to undertake textual authentication and also face recognition. It is not possible to crack the password because the request and response within the server and client won't be having a password in text format. The LSB algorithm is modified using a pseudo-random number generator, which makes the application more robust.

Mudusu et al. [19] proposed the steganographic approach, where the textual information is embedded in the audio of the multimedia source, and the face recognition algorithm is used to improve the security of the data. The multimedia file was first divided into audio and video data. The secrete text message is inserted in the audio data using the LSB method, and face data is embedded in the video file to form a stego multimedia file. The personal key inserted in the file using the RSA calculation.

Nuha Mohammed et al. [20] presented an approach of audio steganography using magic cubes and LSB. This algorithm is based on the mathematical formulation of the magic cube. The value obtained by the magic cube is used as an indexed value of the audio cover. LSB algorithm is used to embed the secrete message in the selected location set by a magic cube. The keys of the magic cube are used to advance the security of the LSB steganography method.

Rashid et al. [21] proposed a high invisibility face biometric transfer technique. For transmission, the extracted face features are embedded with the aid of a robust image steganography method. The invisibility of the system is measured by calculating PSNR between original and stego-image. The recognition process is done in the receiver end after extracting the face features.

Thus we came across a different system that used LSB based audio-video steganography. Therefore, we understand that LSB based steganography is an essential and effective technique for providing data security. While most of the

researches have done it for encrypting the audio and then embedding it back in the audio-video file. The system proposed in this paper goes a step ahead of that the existing works and uses face recognition along with the LSB, which helps in providing additional data security.

3. Proposed System

The proposed system comprises of two modules: Face recognition and encryption followed by decryption. Each module is explained in detail below.

3.1 Face Recognition System

The comprehensive block diagram of the Face Recognition system is demonstrated in Fig. 2.

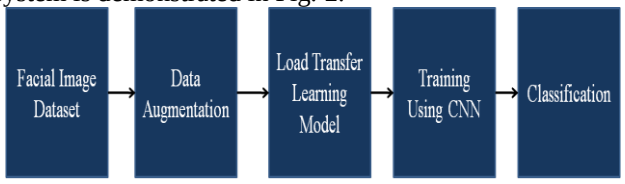


Fig. 2 Block diagram for embedding and retrieval

The main stages are Image Acquisition, Database Creation, Face Detection, Pre-processing, Data Augmentation, Extraction of Content, CNN Training, and Classification. There two main phases: Training phase and Testing (Recognition) phase.

3.1.1 Image Capture

The database is collected in real-time. The database consists of facial images of five persons in different light and luminance conditions with different angles. The collected images are in RGB format of size 227X227 pixels. The database distribution of the training and testing used for this system is tabulated in Table 1.

Table 1: Database distribution for the face recognition system

Data Labels	Total facial Images	Training Facial Images	Testing Facial Images
1	973	779	194
2	830	664	166
3	924	740	184
4	842	674	168
5	1453	1089	364

3.1.2 Pre-processing

Sometimes the captured facial images require a little pre-processing like cropping the face, resizing, histogram equalization for removal of illumination variance, noise reduction, thresholding, converting to the binary, or grayscale image, etc. The input image is in RGB color format. For further processing, the RGB is transformed into a grayscale image. It can be done by combining the three-channel, but this approach fails as the red color has more wavelengths between the three colors, and green color has a lower wavelength than red color and gives eyes a calming effect. Therefore, we recognize that the impact of red color must be decreased, and the effect of green color increased, and the effect of blue color is put between these two. The conversion of RGB to gray is given by

$$0.30 * R + 0.69 * G + 0.11 * B \quad (1)$$

Where, R, G, and B represents the pixel intensity values of the red, green, and a blue pixel, respectively.

3.1.3 Data Augmentation

The practice of artificially creating new training data from previously available training data is referred to as data augmentation. It is done by using domain-specific approaches to generate new and dissimilar training examples from the training data.

Image data augmentation is among the most accepted types of data augmentation. It includes the creation of renovated forms of images within the training dataset that fit into the identical class as that of the original image. The transformation consists of various operations of image manipulation like zooms, flips, shifts, and many more. The goal is to expand the training dataset with new examples possible. Thus the model is likely to observe variations of the training set images.

3.1.4 Training and testing Using CNN

CNN used for image classification problems. It is a multi-layered feed-forward neural network. It consists of neurons with weights, parameters, and biases. CNN's structure includes Convolutional, Pooling, Rectified Linear Unit (ReLU), and Fully Connected Layers (FCL).

3.1.4.1 Convolutional Layer

This layer forms CNN's central building block, which performs most of the heavy computational work. The fundamental principle of the convolution layer is to provide extraction features from the image input data. Using a collection of learnable neurons convolves the input image. It generates a feature map or activation map in the output image, and then the feature maps are fed into the next convolutional layers as input data.

3.1.4.2 Pooling Layer

This layer decreases the dimensionality of each activation map, but it does provide the essential details. The input images are separated into a set of masks that are not overlapping. Every region with a non-linear operation such as average or maximum is down-sampled. This layer reaches

faster convergence, enhanced generalization, stable translation, and modification and is typically embedded in convolutional layers.

3.1.4.3 ReLU Layer

ReLU is a layer that added non-linearity in the data. It is applied per pixel, and all negative values are reconstituted to zero in the feature map. To comprehend how the ReLU operates, we agree that input is given as x, and from that, the rectifier is referred to as f(x)= max(0,x) in the neural network image literature. The aim of using the FCL is to use these functions to classify the input image into different groups based on the training dataset. FCL is measured to be the final pooling layer, which inputs the features to a classifier using the activation function Softmax. The Fully Connected Layer summing up the output possibilities is 1. It is confirmed by making use of the Softmax as the function for activation.

3.1.4.4 AlexNet

AlexNet is a popular deep network used for several computer vision applications. In this approach, the transfer learning of a trained CNN model that is AlexNet is employed for face recognition. The AlexNet model architecture is shown in Fig. 3.

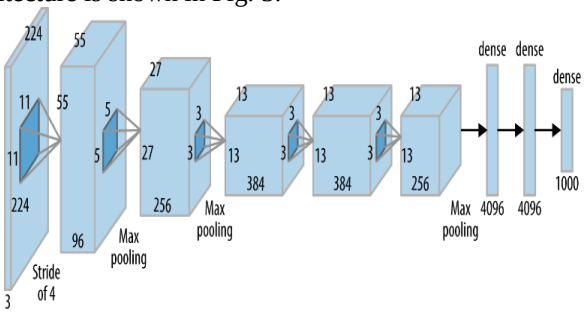


Fig. 3 Architecture of AlexNet

AlexNet has five convolutional layers trailed by three fully connected layers. These convolutional layers extract essential features from the image. Every convolutional layer is composed of linear convolution filters that are trailed by ReLu activation, normalization, and max pooling. The primary layer is the input layer, which takes images having size 227-by-227-by-3. The very first convolution layer has96 filters, each of which is sized 11x11x3 with pace four and no padding. The results of the first convolutional layer are passed to the ReLu layer, which is followed by the max-pooling layer. The purpose behind using the ReLu activation function is the prevention of propagation of any non-positive value in the network. The pooling layer aims to lessen computation and to control overfitting. The second convolutional layer comprises of 256 filters sized 5x5 with pace one and padding 2. The third, fourth, and fifth convolution layer executes 3x3 convolution with speed one and padding 1. Only convolutional layers 1,2, and 5 have max-pooling. Three fully connected layers trail the down-sampling and convolutional layers. The final fully connected layer uses features learned from the prior layer to execute the classification task. This layer is trailed by a softmax layer, which will normalize the output. In this approach, we have trained AlexNet for face recognition. Fig. 4. shows the AlexNet training. Accuracy of 99.66% is achieved during the training.

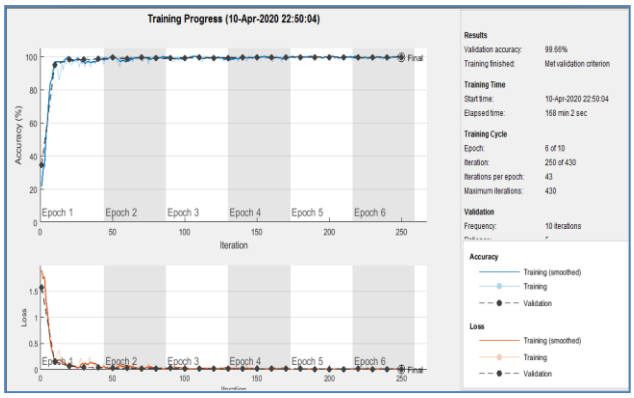


Fig. 4 AlexNet Training Progress

3.2 Audio Steganography (Encryption and decryption)

The block diagram of the audio steganography is as shown in Fig.5. The audio-video file (.avi) is the input for this system. Firstly, the audio from the video multimedia (audio-video) file is extracted. Pseudo-random Noise (PN) sequence is multiply with secret audio referred to as Spread Spectrum (SS) technique and then embedded with the original audio file using the LSB method. This spreading code has a higher chip rate, resulting in a time-continuous scrambled wideband signal. The converted audio data is called a stego-audio file. The video is extracted from the original multimedia file, and the authorized user face image is embedded to generate the stego-video. The LSB algorithm is more straightforward and more widely used because it provides the best audio location to mask the secret audio in such a way that the stego-audio file has minimal distortion. Therefore, the gap between stego-audio and the original audio is smaller. The stego audio and stego-video are again embedded into a video file and transmit over the secure communication channel.

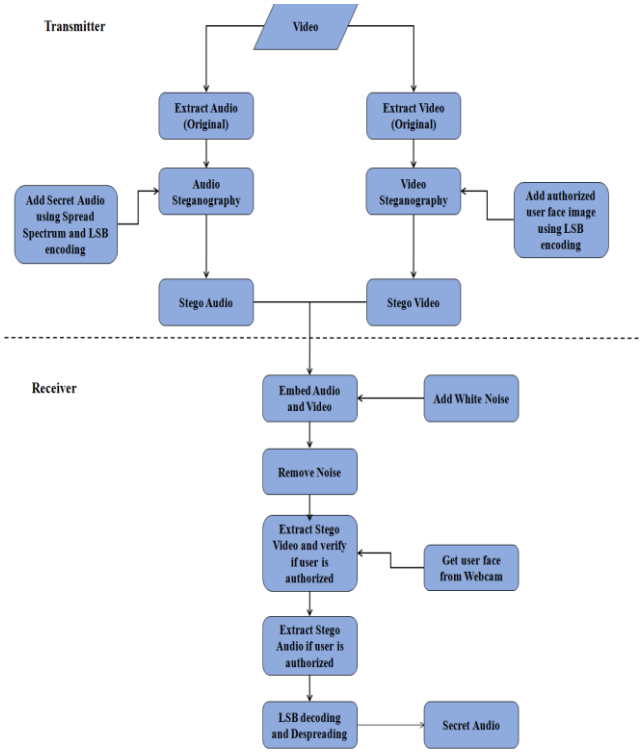


Fig. 5 Block diagram for embedding and retrieval

The obtained video is decrypted in the receiver segment to retrieve the concealed secret message from the audio-video file stego. The video file is extracted from the stego audio-video file in the decryption portion to recover the user image from the chosen frame. Then the input face picture is taken via the webcam and compared to the authentication picture that is covered. When both images suit, then the user will only be able to retrieve the hidden message behind the audio loop and wait until the right recipient appears in front of the webcam. Upon matching the authentication image, the user will extract hidden text from the stego audio-video file. The crucial process of this algorithm is to retrieve the original secret audio from stego-audio. It is executed by using the LSB method and SS technique.

4. Result

This section discussed the performance of the proposed face recognition and LSB based audio steganography method. The proposed system is implemented using MATLAB 2019aX64 bit version.

4.1 Analysis of Face Recognition

The Qualitative analysis is the Subjective judgments of the performance of the system. Fig.6 (a) shows the input from the testing dataset of five authorized persons and one unauthorized person, and the output of the presented face recognition system using CNN is shown in Fig. 6 (b).

(a) (b)

Fig.6 Face recognition (a) Input Image (b) Output recognized image

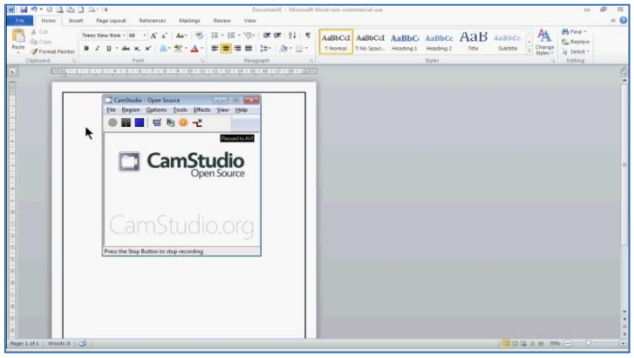
The results of the face recognition show that this CNN based face recognition system able to recognize the face correctly with higher accuracy.

4.2 Analysis LSB based steganography

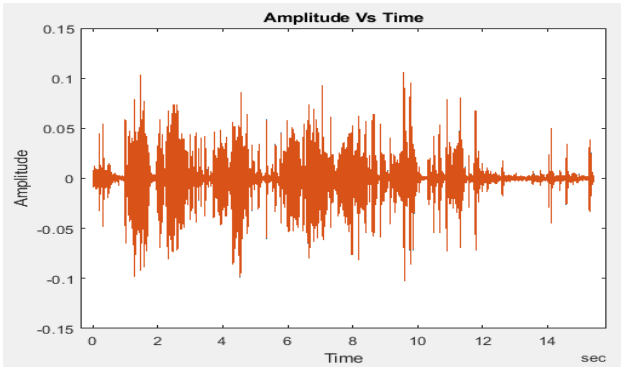
4.2.1 Encryption Process

In this section, visual analysis of the audio-video steganography is explained. Figure 7 (a-g) shows the qualitative analysis of the proposed system. The original audio-video multimedia file (VIDEO_STEGANOGRAPHY.avi) is the input of the system, as shown in Figure 7(a). The audio file is extracted from the audio-video multimedia file called as an original audio file or cover audio file. The waveform of the original audio file is as presented in Figure 7(b). The secrete file (orig_secret.wav) is embedded or inserted in the original audio file using the LSB algorithm. The resultant audio file is called a stego-audio file.

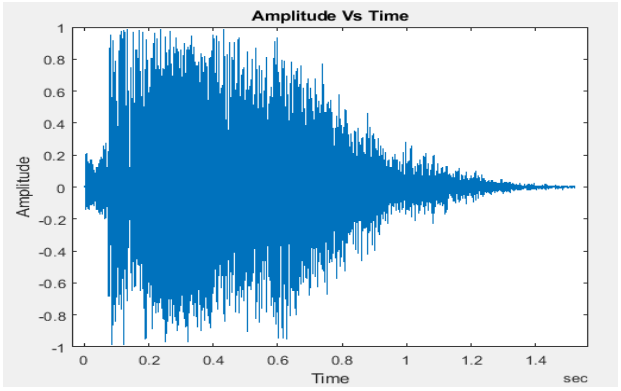
The waveform of the secret audio file is, as shown in Figure 7(c). The secret face file is shown in Figure 7(d), and its binary representative image is shown in Figure 7(e). The final stage-audio file after embedding the secret audio is shown in Figure 7(f). The waveform of the final encrypted stage audio-video file that is to be transmitted over the channel is shown in Figure 7(g).



(a)

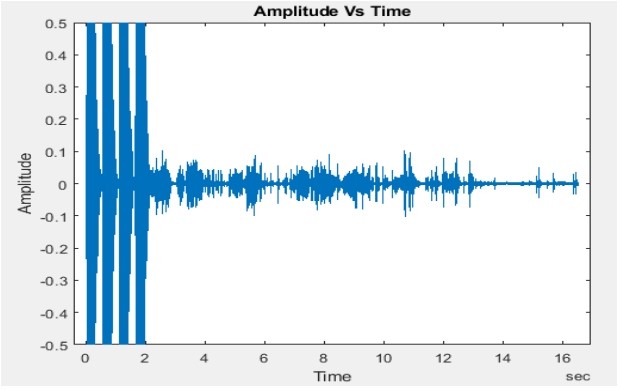


(b)

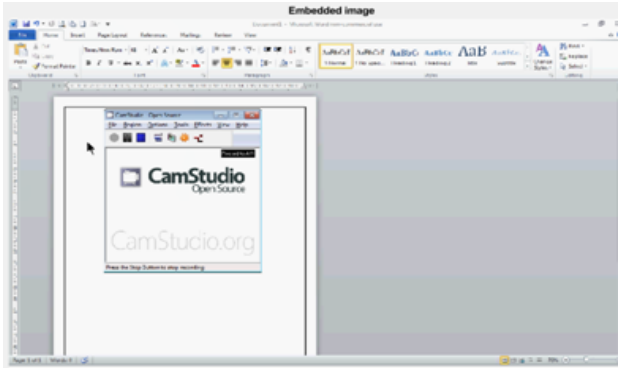


(c)

(d) (e)



(f)



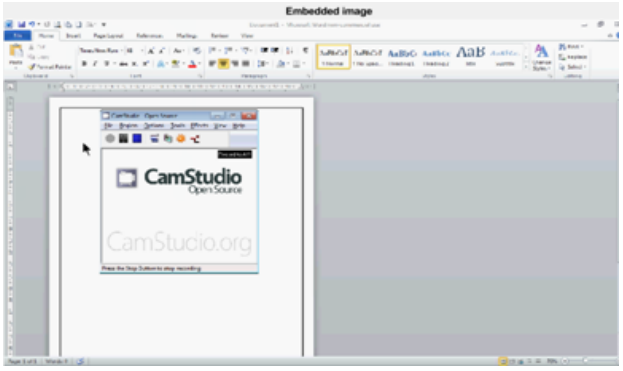
(g)

Fig. 7 Qualitative analysis of the proposed system at transmitter side (a) Input audio-video multimedia file (b) Original audio or cover audio file (c) secret audio file waveform (d) Secrete face file (e) Binarized face file (f) stego audio file (g) stego audio-video-face encrypted file.

According to the graph of cover audio, as shown in Figure 7(b) and the stego audio file, as shown in Figure 7(f). The authentication user image and its binary representation are shown in Figure 7(d-e). These images will be used to match the input image at the receiver side. Finally, the stego audio is embedded in the original video file and transmitted over the communication channel.

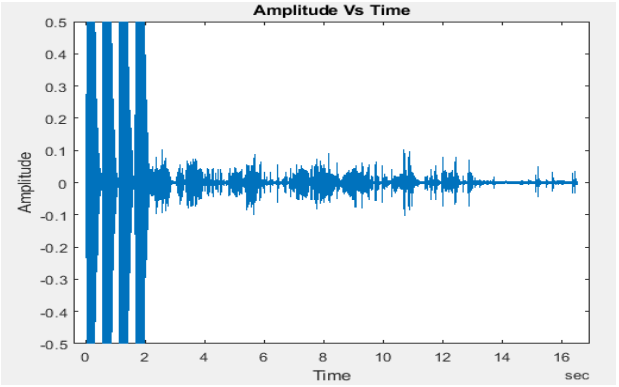
4.2.2 Decryption process

At the receiver side, the stego audio-video file, as shown in Figure 8(a) is received, the recovered binary image is shown in Figure 8(b). Further, the process of extraction of the secret audio from the stego-audio is performed, and the extracted secrete audio is shown in Figure 8(c). The original cover audio is as shown in 8(d).

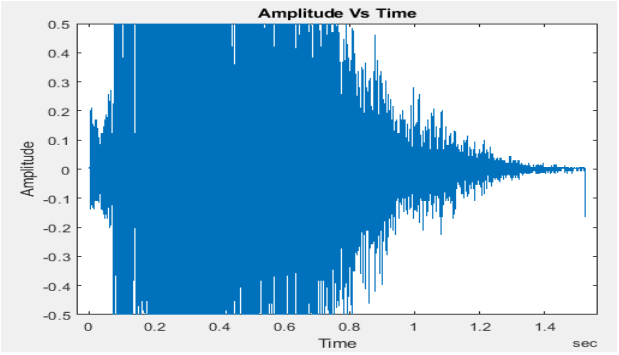


(a)

(b)



(c)



(d)

Fig. 8 Qualitative analysis of the proposed system at the receiver side (a) Stego audio-video-face encrypted file (b) Recovered binary face image (c) Extracted stego audio file (d) Extracted secrete audio file waveform

From the qualitative analysis, it is observed that the recovered binary image is shown in Figure 8(b) is first used to identify the authenticated user by comparing the image at the receiver taken from a webcam with the transmitted image. Once the image matches the extracted stego file is obtained as shown in Figure 8(d), and thus the secret audio file is achieved by final extraction.

In this approach of audio-video steganography, sustainable embedding is introduced that neither alter nor overwrite the bits. Since there is no visual difference between the cover file and stego file, hence no one can predict the presence of stego in the communication channel.

4.3 Quantitative Analysis

The Results of the systems are evaluated using Peak Signal to Noise Ratio (PSNR), Root Mean Square Error (RMSE), and Structural Similarity Index Matrix (SSIM). The detailed explanation of this parameter is as explained as follows.

4.3.1 PSNR

PSNR is the parameter of the audio file that means Peak Signal to Noise Ratio. PSNR and MSE both are inversely proportional to each other, and the following equation can measure PSNR.

$$PSNR = 10log_{10}\left[\frac{I^2}{MSE}\right] \quad (2)$$

Where, I is the maximum possible value of audio.

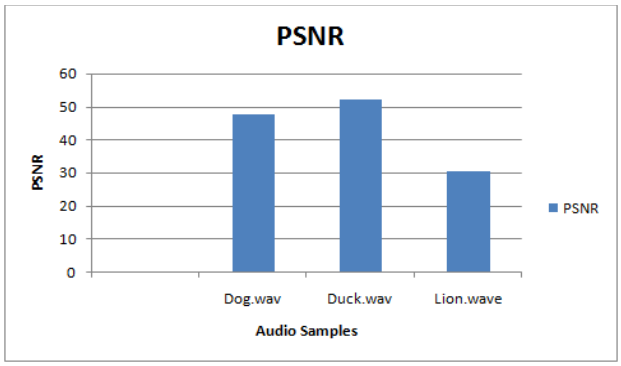


Fig. 9 Graphical analysis of PSNR for different audio samples

4.3.2 RMSE

RMSE is a parameter that means Root Means Square Error which is calculated as the square root of MSE

$$RMSE = \sqrt{\frac{1}{[N\times M]^2}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(X_{ij} - Y_{ij}\right)^2} \quad (3)$$

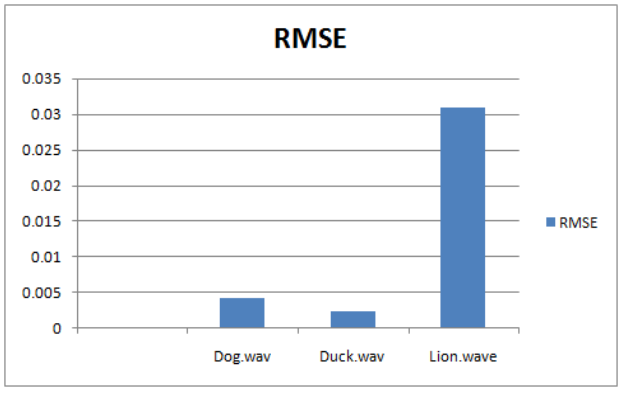


Fig. 10 Graphical analysis of RMSE for different audio samples

4.3.3 SSIM

SSIM is the measure of the quality degradation caused by the modification and loss in the data transmission. The SSIM is calculated in this approach is between the original audio and extracted audio.

$$SSIM(x,y) = \left(2\mu_x\mu_y + C_1\right) \quad (4)$$

where μ_x, μ_y, are the local mean, σ_x, σ_y are the standard deviation and $\sigma_{xy}$is the cross-covariance for data x, y.

The mean, standard deviation, and cross variance is given by

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (5)$$

$$\sigma_x = \left(\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2\right)^{\frac{1}{2}} \quad (6)$$

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y) \quad (7)$$

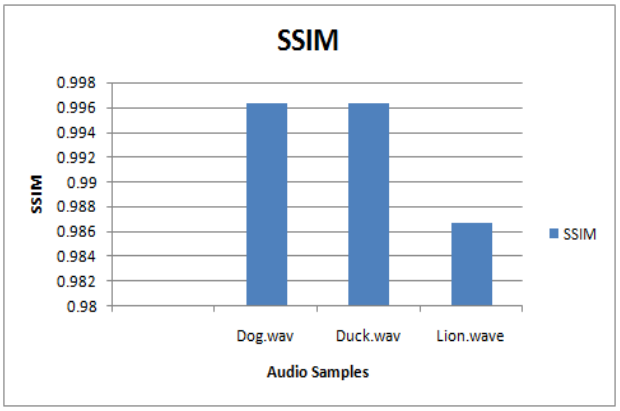


Fig. 11 Graphical analysis of SSIM for different audio samples

From the analysis (Fig. 9, 10, 11), it is understood that the PSNR and SSIM value of the samples is maximum while RMSE is lower. It shows the accurateness of the system.

The robustness of the steganography systems mostly depends on the scaling parameter for the embedding process. Most of the state of art methods used static value of scaling parameters for the simplicity of the algorithm. The work proposed by [23] used the optimal process of selection of scaling factors to obtained high Signal to Noise Ratio. Even the statistic error values like MSE and RMSE have been found in the excellent embedding range. Statistically, there is no difference in the stego and cover file. Hence, it can be concluded that the system is sustainable for different embedding ranges.

The comparative analysis of the proposed system with state of the art methods [22, 23] has been presented in Fig. 12-Fig.14.

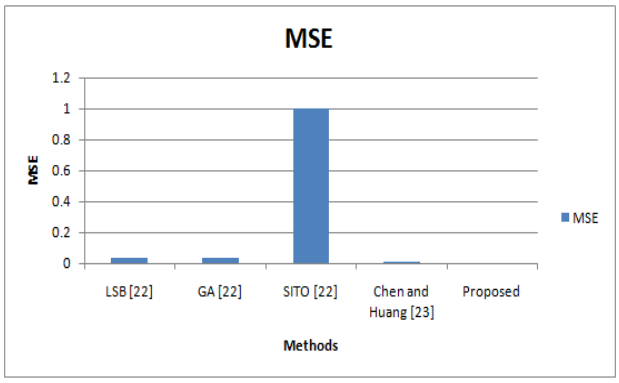


Fig. 12 MSE comparison of the proposed system with existing [22, 23]

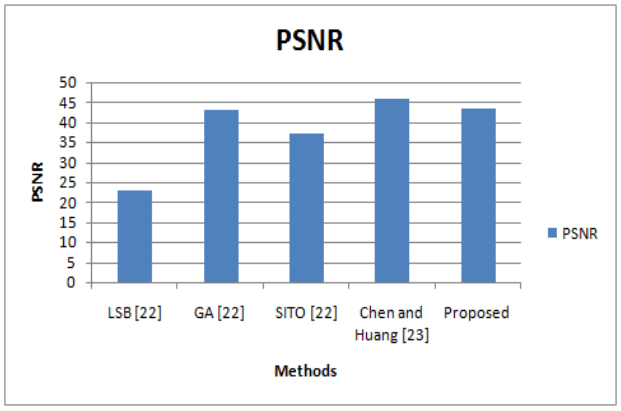


Fig. 13 PSNR comparison of the proposed system with existing [22, 23]

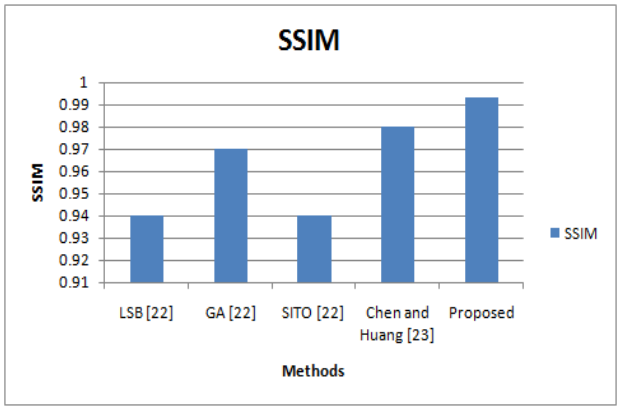


Fig. 14 SSIM comparison of the proposed system with existing [22, 23]

It is observed that from a comparative analysis of MSE, PSNR, and SSIM, the proposed methodology shows better results compare to the state of art methods.

5. Conclusion

In this presented system, an audio-video steganography approach using the Least Significant Bit (LSB) and face recognition has been performed. The proposed method detects stego audio precisely as well as extract the original audio and video frames accurately. It also uses a user image to detect before decrypting the message with the help of the face recognition technique. Thus this system provides enhanced data security than the previous system because it has the advantage of dual protection of face recognition and LSB technique too. The embedding capacity of this system is less than the previous method since we make use of the spread spectrum, and an increase in the secret audio will lead to more spread output, which in turn will increase the transmission bandwidth. Still, it provides the advantages of good robustness and immunity to noise attack as the spread spectrum is by far the most robust data hiding technique. On average, the method embeds a secret audio bit per sample of cover audio. This method could widely use for the modification to LSB's without hampering the audio quality of the sound. The proposed approach attained enhanced MSE, PSNR, and SSIM of 0.000326, 43.3593, and 0.993067, respectively. The proposed method shows superior results than the state of the art method.

References

[1] M. S.Merkow and J. Breithaupt, "Information security: Principles and practices," Pearson Education, 2014.

[2] Esra Satir and Hakan Isik, "A Compression-based text steganography method," Journal of System and Software, Vo. 85, Issue 10, pp.2385-2394, 2012.

[3] Abbas Cheddad, Joan Condell, Kevin Curran, PaulMc Kevitt, "Digital image steganography: Survey and analysis of current methods," Journal of Signal Processing, Vol. 90, Issue 3, pp. 727-752, 2010.

[4] Anastasia Ioannidou, Spyros T. Halkidis, George Stephanides, "A novel technique for image steganography based on high payload method and edge detection," Journal of Expert System with Application, Vol. 39, Issue 14, pp. 11517-11524, 2012.

[5] Mennatallah M. Sadek, Amal S. Khalifa and Mostafa G. M. Mostafa, "Video Steganography: A Comprehensive Review," Journal of Multimedia Tools Applications, Vol. 74, pp. 7063-7094, 2014.

[6] Vaishali Sarangpure, Prof. Roshani Talmale, Prof. G.Rajesh babu, "Implementation on Hiding Data and Image in Audio-Video Using Anti Forensics Technique," Journal of Innovative Research in Computer and Communication Engineering, Vol. 3, Issue 9, pp. 8159-8164, 2015.

[7] Moresh Mukhedkar, Prajkta Powar, Peter Gaikwad, "Secure non real-time image encryption algorithm development using cryptography & steganography," 2015 Annual IEEE India Conference (INDICON), New Delhi, 2015, pp. 1-6.

[8] Ms. Srushti Save, Ms. Pracheta Raut, Ms. Prajakta Jadhav, Ms.Tejaswini Yadav, "Data Security Using Audio- Video Steganography," Journal of Engineering Research & Technology (IJERT), Vol. 7, Issue 2, pp.105-108, 2008.

[9] S. M. Masud Karim, Md. Saifur Rahman, Md. Ismail Hossain, "A new approach for LSB based image steganography using a secret key," 14th International Conference on Computer and Information Technology (ICCIT 2011), Dhaka, Bangladesh, pp. 286-291, 2011.
[10] Nadeem Akhtar, Pragati Johri, Shahbaaz Khan, "Enhancing the Security and Quality of LSB Based Image Steganography," 5th International Conference on Computational Intelligence and Communication Networks, Mathura, pp. 385-390, 2013.
[11] R Kavitha, A Murugan, "Lossless Steganography on AVI File Using Swapping Algorithm" International Conference on Computational Intelligence and Multimedia Applications (ICCIMA), Sivakasi, Tamil Nadu, pp. 83-88, 2007.
[12] Sanjeev Kumar, Amarpal Singh, Manoj Kumar, "Information hiding with adaptive steganography based on novel fuzzy edge identification," Journal of Defence Technology, Vol. 15, Issue 2, pp. 162-169, 2018.
[13] Mukesh Dalal, Manisha Singh, Arun Kumar, Charu, Mamta Juneja, "An Approach of Data Hiding in Video Steganography using Object Detection," Journal of Engineering and Advanced Technology, Vol. 8, Issue 5, pp. 2460-66, 2019.
[14] Mohammad Kamran Khan, Mohammad Naseem, Ibrahim Mohammad Hussain, and Aisha Ajmal, "Distributed Least Significant Bit technique for data hiding in images. Conference on Multitopic Conference, Karachi, pp. 149-154, 2019.
[15] Yugeshwari Kakde, Priyanka Gonnade, Prashant Dahiwale, "Audio-video steganography," IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems, ICIIECS'15, Coimbatore, pp.1-6, 2015.
[16] Najme Maleki, Mehrdad Jalali, Majid Vafaei Jahan, "Adaptive and non-adaptive data hiding methods for grayscale images based on modulus function," Journal of Egyptian Informatics, Vol. 15, Issue 2, pp. 115-127, 2014.
[17] C. Sumanth, M. B.Meenavathi, "Audio-video steganography using face recognition technology for authentication," Journal of Advanced Research in Electronics and Communication Engineering, Vol. 5, Issue 5, 1313-1319, 2016.
[18] Nishant Kaushik, Parveen Sultana H, Senthil Jayavel, "Remote Authentication using Face Recognition with Steganography," International Journal of Recent Technology and Engineering, Vol. 7, Issue 4S, pp. 351-354, 2018.
[19] Rambabu Mudusu, A. Nagesh, M. Sadanandam, "Enhancing Data Security Using Audio-Video Steganography," International Journal of Engineering & Technology, Vol. 7, Issue 2.20, pp. 276-279, 2018.
[20] Nuha Salim Mohammed, Ziyad Tariq Mustafa Al-Tai, "Using Magic Cube and a Modified LSB for Audio Steganography," International Journal of Engineering & Technology, Vol. 7, Issue 4.19, pp. 727-731, 2018.
[21] R. D.Rashid, S. A.Jassim, H. Sellahewa, "Covert exchange of face biometric data using steganography," 5th Computer Science and Electronic Engineering, Colchester, pp. 134-139, 2013.
[22] Rohit Tanwar, Kulvinder Singh, Mazdak Zamani, Amit Verma, and Prashant Kumar, "An Optimized Approach for Secure Data Transmission Using Spread Spectrum Audio Steganography, Chaos Theory, and Social Impact Theory Optimizer," Hindawi Journal of Computer Networks and Communications, pp. 1-10, 2019.
[23] Shuo-Tsung Chen & Huang-Nan Huang, "Optimization-based audio watermarking with integrated quantization embedding," Journal of Multimedia Tools and Applications, Vol. 75, Issue 8, pp. 4735– 4751, 2016.

Alaknanda S. Patil is working as an Assistant Professor in the Department of Electronics and Telecommunication Engineering, JSPM NTC, Pune. She has completed her BE (IE) and ME (Electronics). She has more than 20 publications in reputed journals and conferences She is currently pursuing a Ph.D. degree in Electronics Engineering at Sathyabama Institute of Science and Technology (Deemed to be University) Chennai.

Dr. G. Sundari is working as a Professor in the Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai. She has 23 years of teaching and research experience in the field of wireless sensor networks and image and video processing. She has more than 50 publications in reputed journals and conferences.