Pixel Value Difference based Face Recognition for Mitigation of Secret Message Detection

Alaknanda S. Patil[1], Dr. G. Sundari[2]
Department of ECE
Sathyabama Institute of Science and Technology
Chennai, Tamilnadu, India

***Abstract*—Data security is an important aspect of the modern digital world. Authentication is necessary for the prevention of data from intruders and hackers. Most of the existing system uses textual password which can provide only single-layer security. The textual passwords are simple but they may prone to spyware as well as dictionary attacks. Hence there is a need for a highly secure and multilayer security method. Steganography, the art of hiding the existence of a message by embedding it into another medium, can be exploited in an authentication system. Steganography has emerged as a technology that introduced steganalysis to detect hidden information. In this approach, the multimedia file is the input that is to be transferred over the media. On the transmitter side, the audio and video files are extracted. The secret audio file is embedded with an audio file using the LSB method while the face of the authenticated person is embedded into the video frame using the Pixel Value Differencing (PVD) method. At the receiver side, the face is extracted using the reverse PVD method and authenticated using the Convolutional Neural Network-based face recognition method. After authentication, the secret audio is extracted using the reverse LSB method. The results show that the MSE, RMSE, PSNR, and SSIM of 0.0000045303, 0.0021, 53.5877, and 0.9957, respectively.**

***Keywords*—Audio; Face recognition; Information Security; LSB; Steganography; Video**

I. Introduction

The digital world is evolving rapidly. It means that people are finding new ways of doing old tasks efficiently and creatively. Although it seems a boon, we should not forget that people won't stop using technology to their advantage. It makes the world more vulnerable as it grows. The authentication systems are the ones that require immediate attention [1].

Information hiding methods have been used in different applications for data security. This consists of copyright protection for digital media, watermarking, and steganography. Digital marketing supplies the framework to mark all data object copies with the owner's mark to insert copyright properties. Fingerprinting embeds a distinct signature for each customer purchasing the object [2].

The science of communicating secret data incorporate with communication channels is called Steganography. The communication media will be an image, audio, and video, etc. In the case of a cover image that could be color, grayscale, or even binary in which secret information is embedded, as a result, the stego object is obtained using embedding algorithms [3].

An eavesdropper may decrypt a cryptographic message but he does not even know that a steganographic message exists. Nowadays the issue of illegal copying of music files, books, and software is of critical significance. To solve such a problem steganography is being used, where any information would be encoded in digital media in such a way that it cannot be easily retrieved. There are several forms of steganography depending on the type of medium which is selected as the carrier and these include text, image, audio, video steganography, etc. The main objective of this approach is to provide multi-level security to the audio as well as video data of multimedia files using PVD and LSB algorithms.

This approach is divided into three steps: First, the secret audio sample is embedded into a multimedia file's extracted audio. The Least Significant Bit (LSB) approach is used for audio steganography. Second, the authorized user's facial image is embedded in the multimedia file frame using the Pixel Value Differencing (PVD) method. Third, a face recognition system is used to recognize the face of an authorized user. The database of authorized and unauthorized users is trained using Convolutional Neural Network (CNN) algorithm.

The proposed paper is prepared as follows; Section II offers an overview of audio-video steganography's recent development using different algorithms and their advantages and disadvantages. Section III presents the proposed methodology for the two-stage steganography approach. Section IV demonstrates the results qualitatively and quantitatively. Lastly, the conclusion is given in Section V.

II. Literature Survey

The literature survey of audio and video steganography is described in this section.

G Prasad et al. [4] proposed a method of hiding important information i.e. text, sound, or image which is embedded into an audio cover file using spatial domain techniques. The main aim of this system is to improve the data security of embedded secret audio files. This system uses the LSB technique to embed the secret audio. The performance of the system is evaluated using MSE, PSNR, and SNR. They suggest the further development can be managed to enhance the capacity and improve the robustness of an algorithm.

N. Taneja et al. [5] suggested that the security of the file transfer can be enhanced by combining encoding and digital signature. A digital signature was applied over these files and embedded with the audio file using the LSB technique. The experimental results show the increasing security and content of the hidden textual information. In the future, two encryption algorithms can be used in the file, and steganography applied to increase security.

Sattar B. Sadkhan et al. [6] presented an audio steganography method in which the LSB method is used to hide the data in the audio signal. With few changes, the bit index in stego can be changed reasonably. The method also generates a secret key, as the embedding threshold can hide and retrieve the data. The large amplitude samples produced a high bit index without decreasing the payload availability in the embedding process.

S. M. H. Alwahbani et al. [7] present the audio steganography and encryption approach in which the secret data is embedded into an audio file. Initially, data is encrypted by one-time padding, then the LSB method is applied in the spatial domain to embed the data into an audio file. The experimentation results show the efficiency of the algorithms and the quality of stego sound.

S. E. El-Khamy et al. [8] presented an efficient approach of steganography in the transform domain i.e. Discrete Wavelet Transform (DWT). Initially, the audio sample is spitting into different subbands i.e. detailed and approximate coefficient. Then select the detailed coefficient and repaced by the embedded encrypted image bit thresholded value. The encryption is performed by the RSA method. The pretraiend threshold value help to hide the cipher bits in the detailed component of the audio file. The result and analysis prove the robustness of the system in a noisy environment.

Lindawati et al. [9] presented the encryption method in which the secret message is hidden into an audio file of different formats like MP3 or .wav using the LSB method in the spatial domain. This method can hide .ppt, .docx and .xlsx files. This system is implemented on android phones. The performance of this system is evaluated using PSNR and it shows good PSNR for .wav and MP3 files.

Sattar B. Sadkhan et al. [10], proposed encryption using the AES algorithm. The secret data is encrypted and hide in the spatial domain using the LSB method to ensure confidentiality of the data. This is the simplest but robust technique. The experimental results show that the system is robust for maintaining data confidentiality. The future direction of this research is to encrypt the voice call between two phones.

Y. Bassil [11] suggested the approach to hide the essential data and information like audio samples to the public browser users. Authorized people could use a private browser to access hidden data within the web content. The experiments have provided an excellent way to hide confidential data and ensure that the LSB technique is not found. Important information can be hidden in website videos or image files in the future.

M. Than et al. [12] present the LSB-based data hiding technique. In this approach, the secret message is embedded into the .mp3 file using the LSB method. The traditional LSB method shows the weaken results besides noise. Hence this system proposed compression after combining Echo Hiding techniques.

Kaur N et al. [13] presents the procedures used for Discrete Cosine Transform (DCT) based Steganography and Discrete Wavelet Transform (DWT), and the authors provided various hiding practices or some undisclosed files in image jpg formats. Results were evaluated to know which procedure is good for hiding images.

Islam AUl et al. [14] presented hidden systematic efforts using the major bits of image pixels. The difference between bit number 5 and bit number 6 is measured, and if the outcome is unlike the secret data bit. Then the bit number 5 value is changed. The consequences of this investigation reveal that the projected method advances the signal-to-noise ratio. Several methods and techniques used in stegno scrutiny and spatial representation processes are evaluated. The likely imminent exploration drifts related to steganography safekeeping and substantiation are summarized.

Cheddad et al. [15] presented the skin tone information encryption method in YCbCr colorspace. There is different application which uses YCbCr colorspace such as object detection, video compression. This system first separates each channel of YCbCr and information is hidden in the Cr plane of YCbCr. Hence, part of the skin reviewed to hide the secret message. This system has low embedding capacity because the data is stored in only a single channel.

Kousik Dasgupta et al. [16] developed LSB-based video steganography. Eight hidden information bits are separated in 3,3,2 and embedded in the RGB pixel values of the cover frames respectively. This propagation pattern is taken as blue's chromatic effect on the human eye is stronger than red and green pixels. Video output isn't abandoned, so we could raise the payload. The proposed approach compares with current LSB-based steganography methods, witnessing an encouraging performance.

Sneha Khupse et al. [17] suggested an efficient video steganography system, using ROI instead of the whole frame in a frame. This approach uses human skin tone to mask the message. Morphological dilation and filling are used for skin area identification. Then, the video frames are translated to YCbCr color space, choosing the frame with the least square error for embedding. Hiding the hidden message is achieved inside the Cb portion of the specific video frame. This approach is constrained as only one video frame is considered for the embedding stage.

III. Face Recognition System

The comprehensive block diagram of the Face Recognition system is demonstrated in Fig. 1.

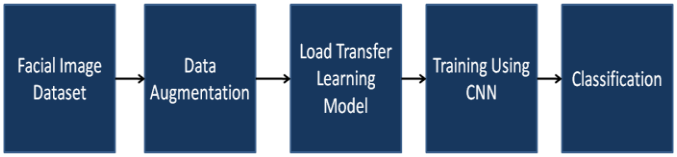


Fig. 1. Block Diagram for Embedding and Retrieval.

The key stages include Image Acquisition, Database development, Face detection, Pre-processing, Data augmentation, Feature extraction, training using CNN, and Classification. There two main phases: The training phase and the Testing (Recognition) phase.

A. 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 I.

B. 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 that can be attained by averaging the three-channel. Still, this method gets failure because Red color has much more wavelengths amongst these three colors, and green color has a lower wavelength than red color and soothes the eyes. Thus, we understand that there is a need to reduce the impact of red color, increase the green color's effect, and impact the blue color within these two [18]. The conversion of RGB to gray is given by.

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

R, G, and B represent the pixel intensity values of red, green, and blue pixels.

C. Data Augmentation

Artificially creating novel data from previously available training data is called data augmentation. Applying domain-specific methods to examples from training data creates new and dissimilar training examples.

Image data increase is among the most accepted data increase types. It includes creating redeveloped image forms within the training dataset that fit in the same class as the original image. The transformation consists of image manipulation operations like zooms, flips, shifts, etc.

The intention is to add new examples to the training dataset. Thus, the model is likely to observe the training set image variations.

TABLE I. 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

D. Training and Testing using CNN

CNN's demonstrated effective image classification. CNN consists of neurons, kernels, or filters with weights, parameters, and biases. Each filter receives inputs, executes convolution. CNN's structure contains Rectified Linear Unit (ReLU) and Fully Connected Layers (FCL).

- Convolutional Layer: The convolutional layer forms CNN's central building block, which performs the heaviest computational work. The primary aim of the convolution layer is to extract input data images. A set of learnable neurons transforms image input. It generates a function map or activation map in the output image and is then fed as input data to the next convolution sheet [19].
- Pooling Layer: The pooling layer reduces the dimensionality of each activation diagram but has the most essential details. Separate input images into non-overlapping rectangles. Each field has an average or maximum non-linear activity. This layer achieves quicker convergence, better generalization, stable translation, and modification, and is usually positioned inside convolutional layers [19].
- ReLU Layer: ReLU is a non-linear operation using the rectifier. Applied per pixel, the map reconstitutes all negative values to 0. To understand how ReLU operates, we accept an input given as x, and in the neural network image literature, the rectifier is called $f(x) = max(0, x)$. Using FCL, these features are used to identify the input image in various groups depending on the training dataset. Using the Softmax activation mechanism, FCL is called the final pooling layer inputting features to a classifier. Summing the maximum layer performance possibilities is 1. Using Softmax as an activation mechanism is verified.

E. AlexNet

AlexNet among the popular deep networks 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. 2.

AlexNet has five convolutional layers trailed by three fully connected layers. These convolutional layers extract essential features from the image. Every convolutional layer comprises linear convolution filters followed 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 has 96 filters, each of which is sized 11x11x3 with pace four and no padding. The first convolutional layer results are passed on 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 control overfitting. The second convolutional layer comprises 256 filters sized 5x5 with pace one and padding 2. The third, fourth, and fifth convolution layer executes 3x3 convolution with rate 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 last layer to execute the classification task. This layer is followed by a softmax layer, which will normalize the output.

In this approach, we have trained AlexNet for face recognition. Fig. 3 shows the AlexNet training. Accuracy of 99.66% is achieved during the training.

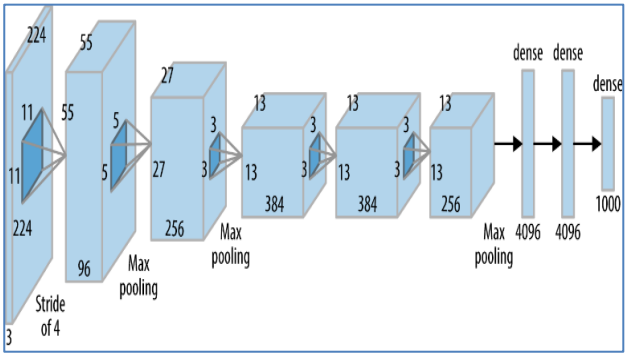


Fig. 2. Architecture of AlexNet.

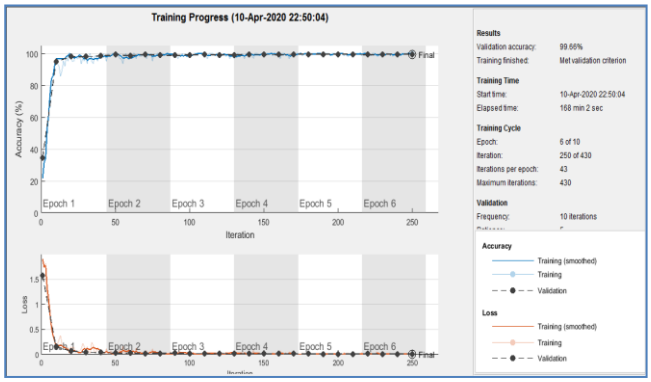


Fig. 3. Alexnet Training Progress.

IV. Proposed System

The proposed audio-video steganography system is as shown in Fig. 4. The proposed audio-video steganography system is divided into two parts transmitter and receiver. In the transmitter section, initially, the audio (called cover audio) and video (cover video) are separated from the multimedia file.

Direct Sequence Spread Spectrum (DSSS) is a spread spectrum technique whereby the secret audio data is multiplied with a pseudorandom noise (PN) spreading code. This spreading code has a higher chip rate/bit rate, which results in a wideband time-continuous scrambled audio. This processed secret audio and cover audio files embedded using the LSB method. If the bit of cover audio C(i,j) is equal to the message bit m of the secret massage to be embedded, C(i,j) will remain unchanged; if not, then set C(i,j) to m. The message embedding processing is as elaborate in Eq. 2 [20].

$$S(i,j) = C(i,j) - 1, if\ LSB\big(C(i,j)\big) = 1\ and\ m = 0$$

$$S(i,j) = C(i,j), if\ LSB\big(C(i,j)\big) = m$$

$$S(i,j) = C(i,j) + 1, if\ LSB\big(C(i,j)\big) = 0\ and\ m = 1 \quad (2)$$

where LSB($C(i,j)$) be the LSB of cover audio, $C(i,j)$and m is been the next message bit which is to be embedded, $S(i,j)$ is the stego audio. The output file is called a stego audio file. The secret audio embedding process is as follows.

1) First, extract the bit from the cover audio.
2) Second, extract the bit from the secret audio
3) Choose the first bit, pick the secret audio and place it in the first component of the bit
4) Place a terminating symbol to indicate the key end. This algorithm used 0 as a terminating symbol.
5) Insert some secret audio file in each first component of the next bit, replacing it
6) Repeat step 6 till all the bit of secret audio has been embedded.
7) Place some terminating symbols to indicate data end.
8) Output stego audio.

In another step, the authorized user's facial image is embedded with the extracted frame of video using the PVD method. The insertion process is explained below.

1) For each sequential pixel ($P_{(i,x)}$and $P_{(i,y)}$) in the cover image, calculate the difference between $P_{(i,x)}$and $P_{(i,y)}$ as d_i. find the lower limit (I_j) and the higher limit (u_j) from the range table (R_j) based on the d_i value.
2) Calculate $w_j = u_j + I_j + 1$
3) Calculate the value of $t_i = \log(w_j)$ with the log base of 2.
4) t_i value determines how many bits can be inserted.
5) Take the t_i message, i_θ is the decimal value of t_i
6) Calculate the value $\delta_i = \theta_i + I_j$
7) Calculate the value of $m = abs(\delta_i - d_i)$
8) Calculate $p'_{(i,x)}$and $p'_{(i,y)}$ by using Eq. 3.

$$p'_{(i,x)}, p'_{(i,y)} = \begin{cases} \left(P_{(i,x)} + \left[\frac{m}{2}\right], P_{(i,y)} - \left[\frac{m}{2}\right]\right), \\ P_{(i,x)} \geq P_{(i,y)} and\ d'_i > d_i; \\ \left(P_{(i,x)} - \left[\frac{m}{2}\right], P_{(i,y)} + \left[\frac{m}{2}\right]\right) \\ P_{(i,x)} < P_{(i,y)} and\ d'_i > d_i; \\ \left(P_{(i,x)} - \left[\frac{m}{2}\right], P_{(i,y)} + \left[\frac{m}{2}\right]\right) \\ P_{(i,x)} \geq P_{(i,y)} and\ d'_i \leq d_i; \\ \left(P_{(i,x)} + \left[\frac{m}{2}\right], P_{(i,y)} - \left[\frac{m}{2}\right]\right) \\ P_{(i,x)} < P_{(i,y)} and\ d'_i \leq d_i; \end{cases} \quad (3)$$

Using the above PVD method, a stego facial image is obtained. The stego audio and stego video are then embedded and transfer over the communication channel with additive white noise.

1) For each successive pixel in the stego image i.e. $p'_{(i,x)}$ and $p'_{(i,y)}$, determine the difference of $p'_{(i,x)}$and $p'_{(i,y)}$ as d'_i. Find the lower limit (I_j) and the higher limit (u_j) from the table range (R_j) based on the value of d'_i,.
2) Determine $w_j = u_j - I_j + 1$.
3) Calculate the value of $t_i = \log(w_j)$ with log_2.
4) t_i the value determines how many bits can be inserted.

5) Calculated $d_i^{"} = d_i' - l_j$, convert $d_i^{"}$ into binary values with the length of t_i

6) The conversion of $d_i^{"}$ into binary with the length of t_i is the hidden message.

7) Differential value algorithm of 2 pixels: finding a separate value of two adjacent pixels, two-pixel width difference, converting d to binary with t length, bit message length,

message power inserted in t in a bit, converting inserted messages to decimal, measuring two new pixels after inserting messages.

On the receiver side, the white noise from the stego multimedia file is removed. The decryption process is started with the separation of stego audio and stego video from the noiseless multimedia file. First, the face from the stego video file is extracted using the reverse PVD method. The process is explained.

The transmitted face is extracted from the above reverse PVD process, which is the face recognition process. The CNN algorithm is used to recognize the face of an authorized user. The extracted face is tested with a trained face recognition model, which gives us whether the user is authorized. The secret audio from stego audio proceeds if the user is authorized. The secret audio from stego audio is extracted using the reverse LSB method. The extraction process is as follows.

1) Extract the bit of the stego audio.

2) Now, start from the first bit and extract the stego bit from the first component.

3) Repeat step 2 till all the bit of secret audio has been extracted.

4) Obtained secret audio.

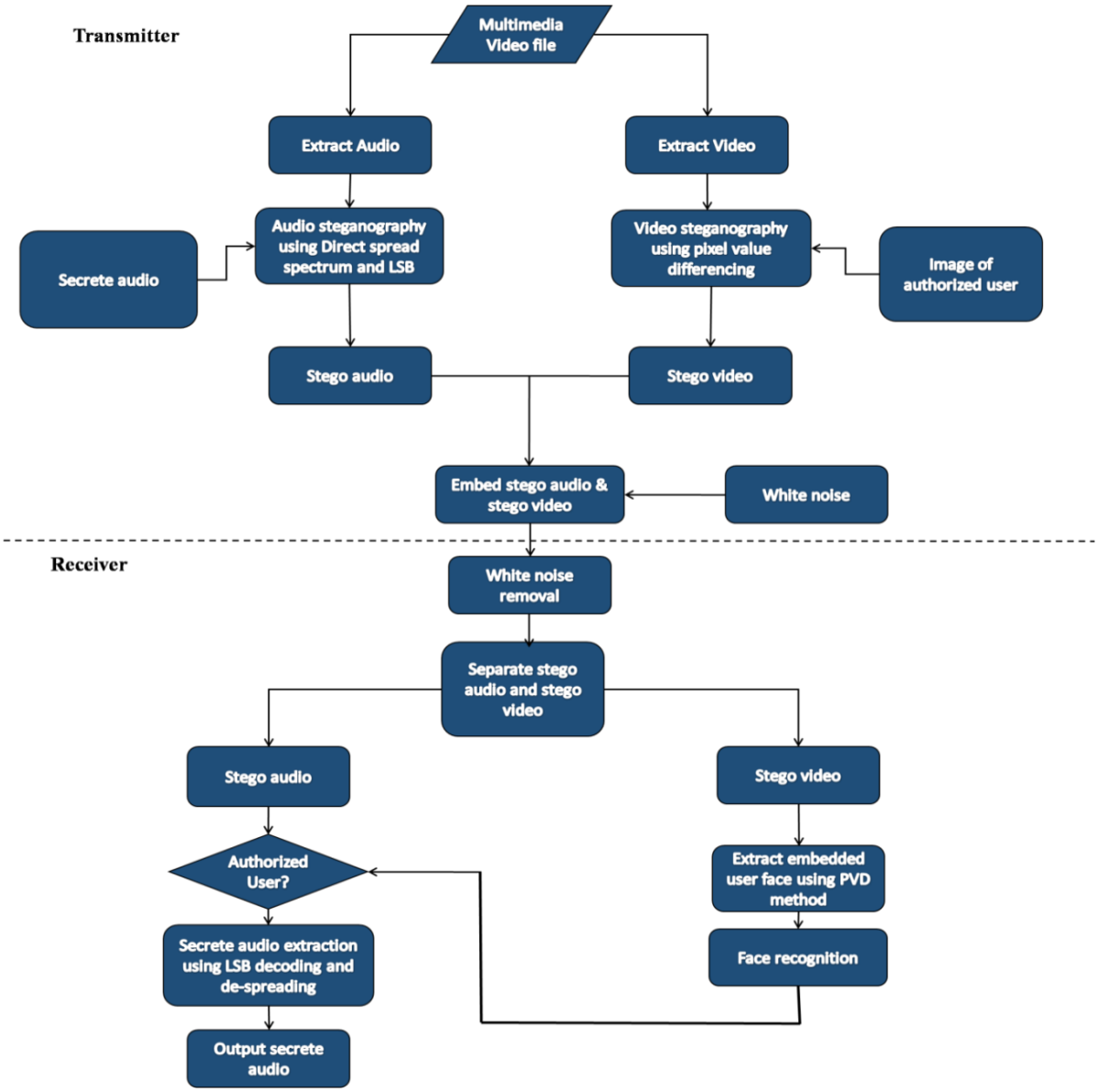


Fig. 4. Block Diagram of Proposed Audio Video Steganography.

V. RESULT

The proposed system is implemented using MATLAB 2020a, the X64 bit software.

A. Analysis of Face Recognition

In this approach, the face recognition system is implemented using a deep CNN algorithm. The training parameter used to train the face recognition system is as tabulated in Table II.

TABLE II. TRAINING PARAMETER OF CNN ALGORITHM FOR THE PROPOSED FACE RECOGNITION SYSTEM

Training Parameters	**Values**
Training algorithm	*'sgdm'*
Momentum	0.9000
Batch size	10
Initial Learning Rate	3e-4
Drop Period	10
Drop Factor	0.1
Gradient Threshold Method	'l2norm'

The face recognition system is implemented using the CNN algorithm.

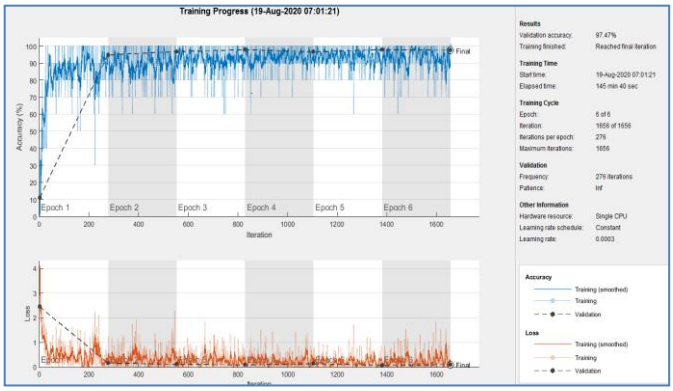


Fig. 5. Traning Progress Graph of CNN based Face Recognition System.

The training progress graph given in Fig. 5 shows that training and validation accuracy graphs follow each other and achieved a validation accuracy of 97.47%. The trained model is portable and used to recognize the authorized and unauthorized user with high accuracy. The qualitative analysis of the face recognition system is, as shown in Fig. 6.

The proposed system's qualitative analysis shows that the proposed system accurately classified the facial samples into authorized and unauthorized users. Fig. 6(a-d) are the authorized samples, while Fig. 6(e) is the unauthorized user's sample face, which is provided to the trained CNN model's input. The CNN model gives accurate output for each sample.

B. Analysis PVD based Video Steganography

In this method, the video frame is considered the cover image, and the face image of the authentic user is regarded as a secret image. The video steganography aims to hide the video frame's facial image, which is not visible to the intruder. The results of the PVD-based video steganography are as shown in Fig. 7.

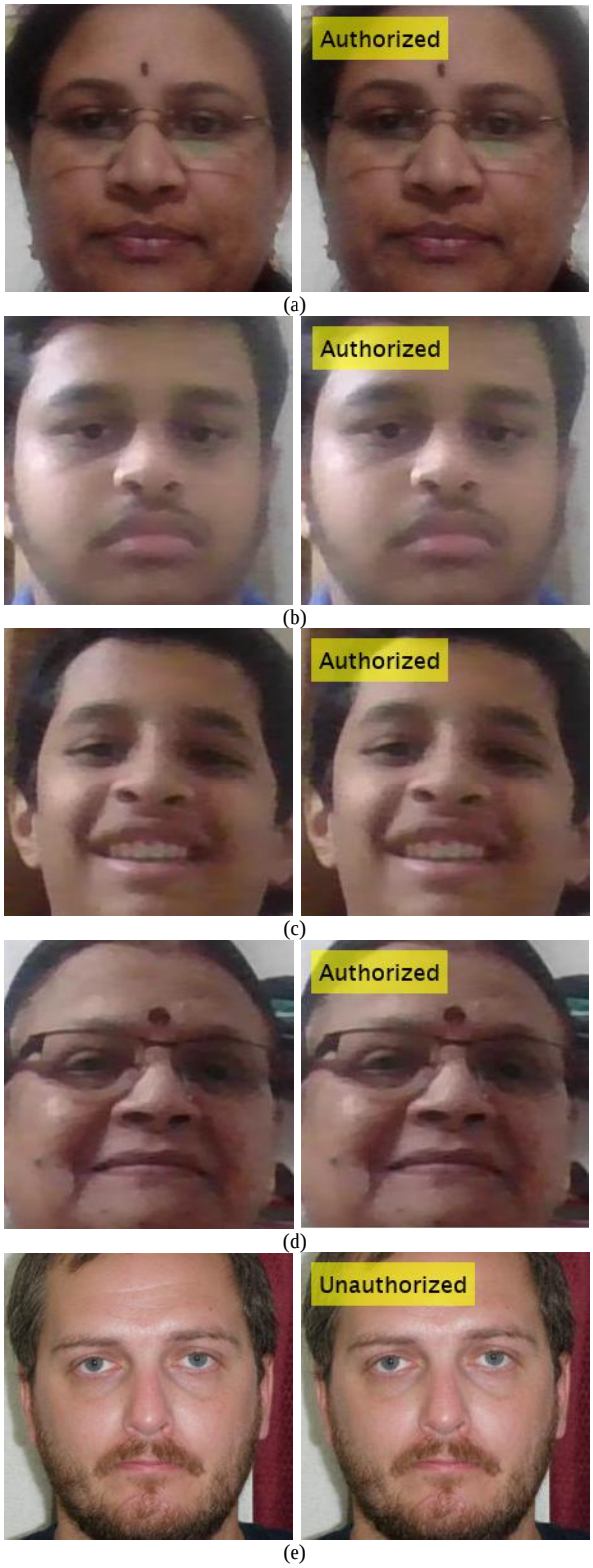
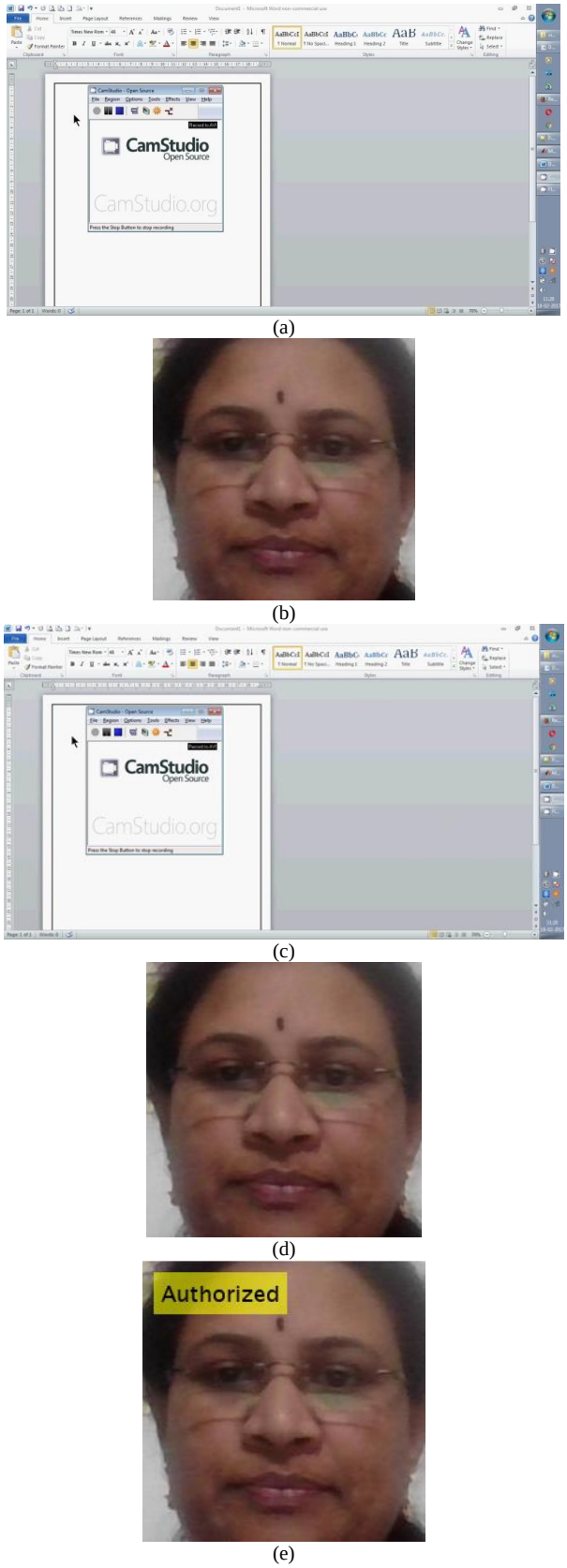


Fig. 6. Qualitative Analysis of Face Recognition System (a)-(d) Authorized User (e) Unauthorized user.



(a) (b) (c) (d) (e)

Fig. 7. Results of PVD based Video Steganography (a) Cover Frame (b) Secret Image (c) Stego Frame (d) Decrypted Secret Image (e) Authorized Face Recognition Output.

Fig. 7 shows the results of the PVD-based video steganography. The cover image and secret image are shown in Fig. 7(a) and Fig. 7(b). The authorized user's secret image is embedded in the cover image using the PVD method and created a stego image shown in Fig. 7(c). Fig. 7(c) shows that the cover image and stego image are visually similar. Hence the intruder cannot predict the embedded secret message with naked eyes. The reverse PVD method is used at the receiver side to extract the face of the authorized user, shown in Fig. 7(d). Finally, the extracted face is tested with a CNN-trained model to predict the authorized person, shown in Fig. 7(e).

C. Analysis LSB based Audio Steganography

In this section, the LSB-based audio steganography process is presented with the analysis shown in Fig. 8. The secret audio and the extracted cover audio extracted from the video multimedia file are presented in Fig. 8(a) and Fig. 8(b). The secret audio and cover audio are embedded using the LSB method called stego audio, presented in Fig. 8(c). From waveform analysis, it is observed that the cover audio waveform and stego audio waveform are looking visually similar. Therefore, it is a problematic intruder to recognize the embedded secret audio. In the decrypted process, the stego audio is extracted using the reverse LSB method shown in Fig. 8(d).

D. 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.

- 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.

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

Where I is the maximum possible value of audio.

- RMSE: RMSE is a parameter that means Root Means Square Error, calculated as the square root of MSE.

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

- 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.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}C+_2)}{(\mu_x+\mu_y+C_1)(\sigma_x+\sigma_y+C_2)} \quad (4)$$

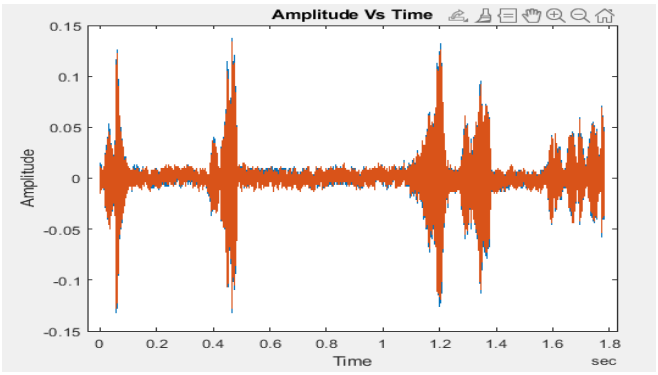
where μ_x, μ_y, are the local mean, σ_x, σ_y are the standard deviation and σ_{xy} is the cross-covariance for data x, y.

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

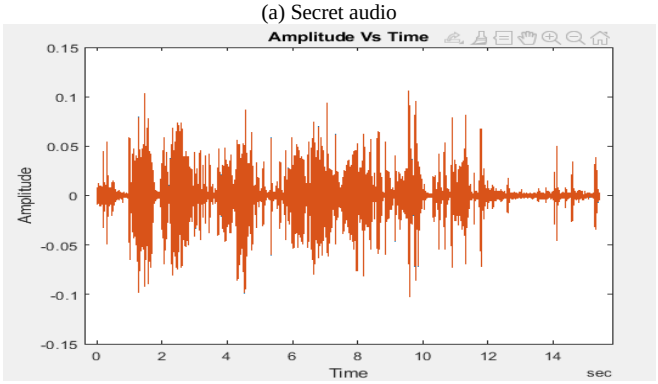
$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (5)$$

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

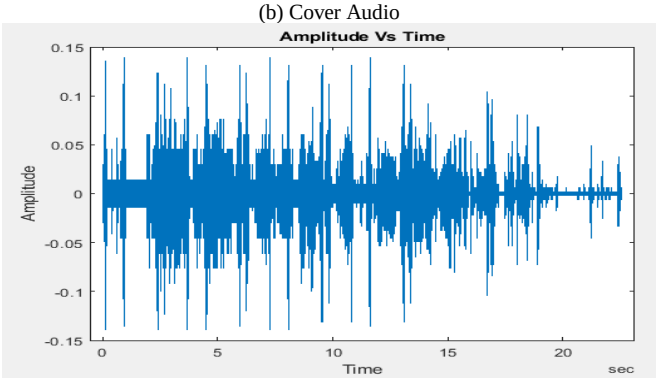
$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y) \quad (7)$$



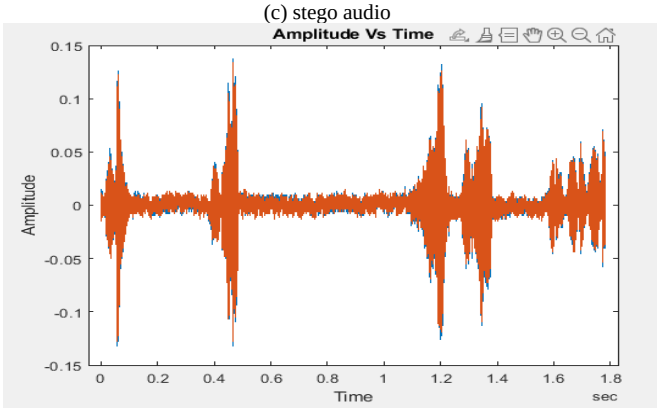
(a) Secret audio



(b) Cover Audio



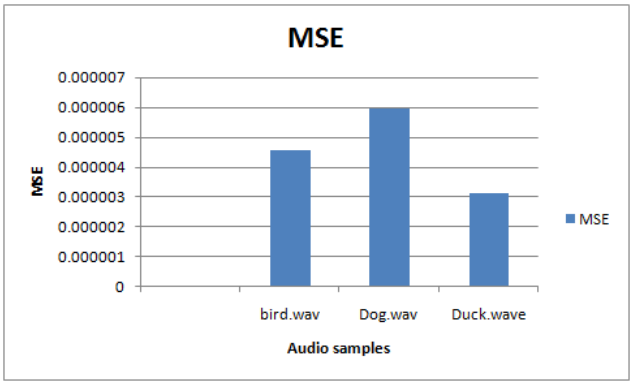
(c) stego audio



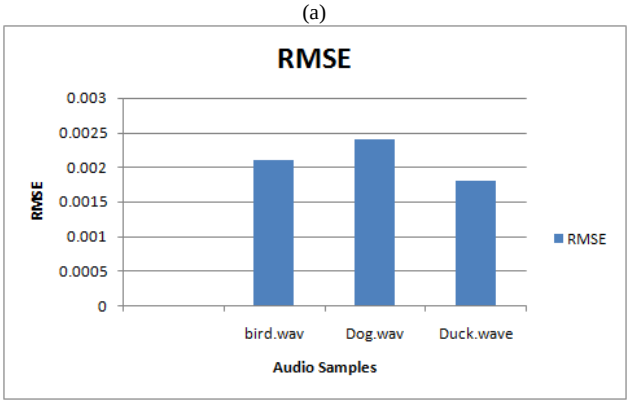
(d)

Fig. 8. Results of LSB based Audio Steganography (a) Secret Audio (b) Cover Audio (c) Stego Audio (d) Decrypted Secret Audio.

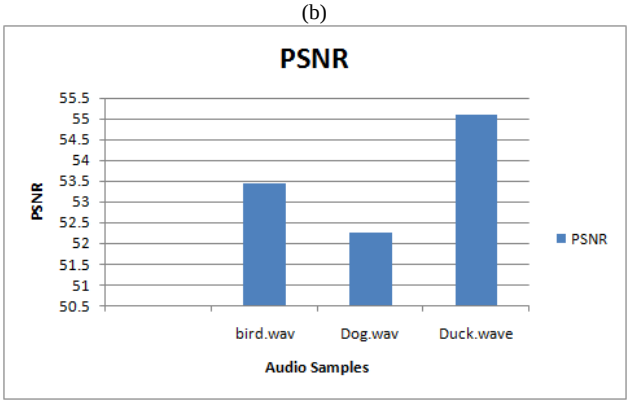
The qualitative analysis in terms of MSE, RMSE, PSNR, and SSIM of the audio steganography is shown in Fig. 9(a-d).



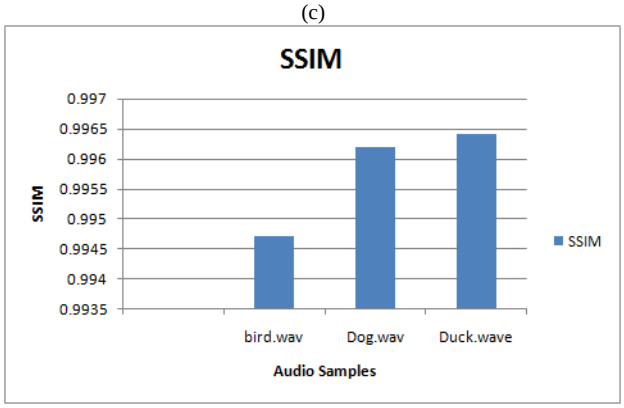
(a)



(b)



(c)



(d)

Fig. 9. Graphical Analysis of the Audio Steganography (a) MSE (b) RMSE (c) PSNR (d) SSIM.

The graphical analysis shows that the MSE and RMSE values of the original secret audio and extracted secret audio are minimal. In contrast, PSNR and SSIM values are high. Hence it can be concluded that the system is highly precise and can be used for steganographic applications.

VI. Conclusion

In this paper, the Audio and video steganography approach is presented. The video steganography is performed using a pixel value differencing method while audio steganography is performed using the least significant method. The authorized user is recognized using the CNN algorithm, which shows excellent validation accuracy of 97.47%. The Results of the system are presented using MSE, RMSE, PSNR, and SSIM. The results show that audio and video steganography leads to promising results. This method could widely be used to modify LSB's without hampering the audio quality of the sound. The proposed approach attained enhanced MSE, RMSE, PSNR, and SSIM of 0.0000045303, 0.0021, 53.5877, and 0.9957, respectively.

In the future, new data-hiding schemes will be worked to improve the embedding capacity by merging the PVD scheme and hidden sharing scheme.

References

[1] N. Kaushik, P. Sultana H, S. Jayavel, "Remote Authentication using Face Recognition with Steganography", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7 Issue-4S, November 2018, pp. 351-354.

[2] Yusuf Perwej, Firoj Parwej, Asif Perwej, "An Adaptive Watermarking Technique for the copyright of digital images and Digital Image Protection", The International Journal of Multimedia & Its Applications (IJMA), April 2012, Volume 4, Number 2, Pp. 21-38.

[3] Nagham Hamid, Abid Yahya, R. Badlishah Ahmad & Osamah M. Al-Qershi, "Image Steganography Techniques: An Overview", International Journal of Computer Science and Security (IJCSS), Volume (6): Issue (3): 2012 168-187.

[4] G. Prasad TVS and S. Varadarajan, "A Novel Hybrid Audio Steganography for Imperceptible Data Hiding," IEEE 978-1-4799-8081-9/15/$31.00 ©, 2015.

[5] N. Taneja and P. Gupta, "Implementation of Dual Security through DSA and Audio Steganography," International Conference on Green Computing and Internet of Things (ICGCIoT), Noida, India, 2015.

[6] Sattar B. Sadkhan, Dr. Nidaa A. Abbas, "Multidisciplinary Perspectives in Cryptology and Information Security", Book, Publisher IGI Global, 2014.

[7] S. M.H. Alwahbani and H. T.I. Elshoush, "Chaos-Based Audio Steganography and Cryptography Using LSB Method and One-Time Pad," Springer International Publishing AG 2018 Y. Bi et al. (eds.), Proceedings of SAI Intelligent Systems Conference (IntelliSys), 2016.

[8] S. E. El-Khamy, N. O. Korany, and M. H. El-Sherif, "A security-enhanced robust audio steganography algorithm for image hiding using sample comparison in the discrete wavelet transform domain and RSA encryption," Multimed Tools Appl # Springer Science+Business Media New York, 2016.

[9] Lindawati and R. Siburian, "Steganography Implementation on Android Smartphone Using the LSB (Least Significant Bit) to MP3 and WAV Audio," IEEE The 3rd International Conference on Wireless and Telematics 2017.Jule 27-28, Palembang Indonesia. 2017.

[10] Sattar B. Sadkhan; Akbal O. Salman, "A survey on lightweight-cryptography status and future challenges2018 International Conference on Advance of Sustainable Engineering and its Application (ICASEA).

[11] Y.Bassil, "Audio Steganography Method for Building the Deep Web," American Journal Engineering Research (AJER) e-ISSN: 2320-0847 ISSN: 2320-0936 Volume-8, Issue-5, 2019, pp-45-51.

[12] M. Than and S. Sin, "Secure Data Transmission in MP3 file using LSB and Echo Hiding," International Journal of Advanced Research in Computer Science, 10(4), 45, 2019.

[13] Kaur N, Bansal A (2014) A review on Digital image Steganography (JCST). /International Journal of Computer Science and Information Technology 5:8135-8137.

[14] Islam AUl, Khalid F, Shah M, Khan Z, Toqeer Mahmood, et al. (2016) An improved image Steganography Technique based on MSB using Bit Differencing, IEEE 978-98.

[15] Cheddad A, Curran K, Condell J, McKevitt P, "Skin tone based Steganography in video files exploiting the YCbCr color space", IEEE International Conference on Multimedia and Expo, 2008, pages 905–908.

[16] Kousik Dasgupta, J.K.Mandal, Paramartha Dutta, "Hash Based Least Significant Bit Technique for Video Steganography(HLSB)," in International Journal of Security, Privacy and Trust Management (IJSPTM), pp 1-11, April 2012.

[17] S Khupse, N Patil, "An Adaptive Steganography Technique for Videos Using Steganoflage", International Conference on Information and Computer Technologies pages 811-815, 2014.

[18] Alaknanda S. Patil1 and Dr. G. Sundari, "Enhancing Data Security in video Steganography using Face Recognition", IJCSNS International Journal of Computer Science and Network Security, VOL.20 No.8, August 2020, pp. 134-144.

[19] M. Coşkun, A. Uçar, Ö. Yildirim and Y. Demir, "Face recognition based on convolutional neural network," 2017 International Conference on Modern Electrical and Energy Systems (MEES), Kremenchuk, Ukraine, 2017, pp. 376-379.

[20] Deb Sunder Swami, Kandarpa Kumar Sarma, "chapter 8 A logistic-Map-Based PN Sequence for Stochastic Wireless Channels", IGI Global, 2017.