International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Image Steganography using Deep Vision Transformer with Piecewise Spread Spectrum for Secure Communication

V. Arunkumar[1], Dr. K. Padmanabhan[2]

[1]Ph.D Research Scholar, Department of Computer Science, Periyar University, Salem, India
Email: *veere.arun[at]gmail.com*

[2]Controller of Examinations, Vivekanandha College of Arts and Sciences for Women (Autonomous), Tiruchengode, Namakkal, TN, India
Email: *padmaindia75[at]gmail.com*

Abstract: ***Image steganography strengthens secure communication by hiding the existence of a hidden message, making certain that only the sender and intended recipient are well informed of it. By embedding encrypted data within the image, an extra layer of security is said to be provided, making it laborious for unauthorized users to access hidden information. Existing image steganography algorithm has experienced large trade-off between payload capacity, image quality and security. In addition, image steganography algorithm lacks the adaptability to maintain the reversibility and high visual quality. In order to address these issues, quality enhanced image stegnographic method with spread spectrum called, Functional Segmentation Deep Vision Transformer with Piecewise Spread Spectrum (FSDVT-PSS) for secured communication is proposed. The FSDVT-PSS method uses Deep Vision Transformer Spread Spectrum Analysis for secured image transmission. Initially to de-noise input medical images for obtaining quality enhanced results Guided Filter Image Pre-processing is applied. After that, pre-processed sample image is split into fixed-size patches to generate patch embedding results. With the aid of Vision Transformer (ViT) cover image is analyzed and then optimal locations are identified for embedding spread spectrum modulated data. The results are stored in Patch Embedding Vector. Next, Positional Encoding is employed to retain spatial arrangement with results stored in Positional Encoding Vector. The results of Patch Embedding Vector and Positional Encoding Vector are combined to model Transformer Encoder layer, therefore generating quality enhanced results. Finally a non-linear activation function is applied as classifier in the classification head to ensure secure communication between the intended sender and recipient via quality enhanced images. The performance of the FSDVT-PSS method is evaluated using different evaluation metrics. The obtained results show that our FSDVT-PSS method outperforms existing ones, with high data confidentiality, data integrity and a lower training time. According to experimental results, this FSDVT-PSS method outperforms most previous works in terms of accuracy and PSNR.***

Keywords: Image steganography, Guided Filter, Functional Segmentation, Deep Vision Transformer, Spread Spectrum

1. Introduction

Secured data transmission is an essential one for improving data confidentiality during secured transmission. Steganography embed the secret information in digital images to protect the sensitive data from unauthorized access. Also quality enhanced image steganography plays a major role in secure communication

A novel reversible data hiding (RDH) method was proposed in [1] that used four stego images to improve embedding capacity, visual quality and security. The method employing Huffman coding reduced perceptual distortion while preserving structural similarity between cover and stego, therefore achieving high PSNR values. Moreover, by employing encryption, the embedded data was further secured by maximizing payload capacity, therefore making certain inaccessible in the absence of both decryption key and four stego images, therefore improving structural similarity index measure (SSIM) with minimal mean square error (MSE).

An invertible neural network (INN) based image steganography was proposed in [2] to ensure secure lossless image steganography. Initially, to achieve better invisibility, an invertible hiding model was applied to avoid information loss, thereby solving the issues arising from ill-posed images. Following which to address issues arising from high capacity, learning cost during image embedding were reduced by only fitting part of color channels. Finally, the conception of a key to constrain embedding process of secret information was also included that in turn enhanced the hidden data security.

Most prevailing steganographic techniques are specifically designed for grayscale images. However, when applied to color images without taking into consideration the inter-channel color interactions, security aspect is said to be compromised. In [3], a novel color image steganography method employing generative adversarial net works was proposed to ensure security for color images. Conventional Linguistic Steganographic mechanisms lack automation. To focus on this aspect Recurrent Neural Networks (RNN) Steganographic methods was designed in [4] ensuring security. Motivated by the fact that human eyes perceive pixel perturbation differently, a novel Deeply-Recursive Attention Network was proposed in [5] to ascertain pertinent areas for information hiding with improved accuracy.

With evolution of computer and communication technology, extensive amount of images are found to be stored in cloud and transmitted via the internet. However, mechanisms to keep sensitive images from being accessed by unauthorized persons still remain unaddressed.

A novel method employing Transformer mechanism for extracting informative feature in steganography was designed in [6]. Also to improve security of secret images, an image encryption algorithm employing recursive permutation was also designed. Yet another method employing deep learning was proposed in [7] to focus on the noise aspects. An enhanced method employing multi-layered data embedding was designed in [8] to ensure secure and reliable solution.

Steganography, the utilization of algorithms for embedding secret information is extensively employed in area of information transmission. However, steganalysis tools constructed employing conventional steganographic mechanism can secret message effortlessly. To ensure secure communication Wasserstein distance along with weight clipping was introduced in [9] to not only boost information capacity but also improve extraction accuracy. A review of image steganography algorithm along with their applications in the domain of military was investigated in [10]. Nevertheless, this method included high variability. To address on this aspect, a data hiding mechanism was designed in [11] employing flipping mechanism. By using this mechanism not only minimized variability but also reduced time complexity extensively.

A framework to ensure security and reduce resource consumption employing lightweight steganography detection was proposed in [12]. Here, multiple residual structures along with Transformer (ResFormer) were used to ensure higher detection accuracy. Yet another GAN based image steganography was presented in [13] to ensure secure communication. Nevertheless, image-based steganography is a demanding task owing to security, computational efficiency and payload. A least significant bit substitution was employed in [14] to not only ensure security but also achieve higher PSNR results. A hybrid steganography method was designed in [15] showcasing superior visual image quality. The proposed mechanism provided robust solution necessitating high levels of both data integrity and security.

1.1 Research gap

The main objective of image steganographic quality enhanced secure communication for medical images remains in initially obtaining separate quality enhanced images via vision transformer and then ensuring security aspects via spread spectrum technology to ensure robust data confidentiality and data integrity thus improving training time with improved accuracy. Using Reversible Data Hiding [1] though ensured SSIM with minimal MSE, however, the training time and accuracy aspects were not focused. Invertible neural network (INN) though with the aid of invertible hiding model resulted in enhanced hidden data security however risk of data confidentiality and data integrity was not concentrated during the simulation process that remains to be one of the major factors for image steganographic quality enhanced secure communication.

1.2 Contributions of the work

Motivated by the above issues, like, data confidentiality, data integrity, accuracy and training time for image steganographic quality enhanced secure communication, in this work, a hybrid method called Functional Segmentation Deep Vision Transformer with Piecewise Spread Spectrum (FSDVT-PSS) with spread spectrum is proposed. The major contributions of this work are pointed below.

- To present a significant method for steganographic quality enhanced secure communication involved in medical images called, Functional Segmentation Deep Vision Transformer with Piecewise Spread Spectrum (FSDVT-PSS).
- To improve training time and accuracy using Piecewise Smooth Approximate Functional Segmentation model is designed for CT medical images.
- To propose a Vision Transformer integrated with Spread Spectrum for quality enhanced secure communication via multi-head self-attention mechanism with improved data confidentiality.
- To embed spread spectrum by spreading secret data across wider area rather than employing pseudo random sequence employing Renyi entropy function, therefore improving data integrity.
- Finally, performance of the proposed FSDVT-PSS method is compared with the conventional state-of-the-art methods.

1.3 Organization of the work

The rest of the paper is organized as given below. Section 2 provides the related works on the image steganographic quality enhanced secure communication using vision transformer techniques. Section 3 displays a brief description of the method called, Functional Segmentation Deep Vision Transformer with Piecewise Spread Spectrum (FSDVT-PSS). After that, Section 4 introduces the qualitative analysis. Following which Section 5 introduces a comprehensive evaluation study between the proposed FSDVT-PSS method and two other existing methods using table and graphical representation. Finally, Section 6 concludes the paper.

2. Related Works

In this cyber era, transmission of data within numerous digital media via public networks plays a major role in covert communication. As far as public network is concerned, communicating sides must make certain security to keep data private and confidential. In spite of the prevailing CNN based methods generate good results, however, performance issues like classification accuracy and stability in network training are said to be compromised.

A novel method with CNN architecture to boost hidden data detection accuracy and stability in spatial domain images was presented in [16]. Owing to irreversibility of process of producing secret images from message codeword, recovery accuracy is very poor. To focus on this aspect, a robust joint coverless image steganography method was proposed in [17] guaranteeing robustness against various attacks. A review of image steganography methods for secure communication was investigated in [18]. Nevertheless, prevailing coverless image steganography methods frequently necessitate both sender and recipient to transmit additional information along with image blocks' locations that will raise serious suspicion. To address on this issue, a robust coverless image steganography based on Speeded-Up Robust Features (SURF) was proposed in [19]. Bio-inspire algorithms were applied in [20] to hide greater volume of information achieving security with efficient data hiding.

Steganography improves data security by embedding within digital channels. In spite of the hospital systems concentrate on both text and medical images, integration of audio still remains untapped potential. The security lapse was addressed in [21] gap by exploring audio data embedding in medical images to improve protection and information integrity. Yet another method to ensure security of message being delivered hiding capacity optimization along with layered message security was presented in [22].

Image steganography, a mechanism of embedding hidden information in digital photographs, should achieve both maximum data hiding and cover media integrity. However, the contemporary steganography methods are at best a compromise between data hiding and cover media integrity. In [23], a method called, Ant Colony Optimization (ACO) Least Significant Bit (LSB) was proposed with the intent of optimizing steganographic embedding capacity. A security mechanism employing spread spectrum was designed in [24] for improving steganographic detection. Yet another method to improve sensitivity by fine tuning hyper parameter was presented in [25] with improved detection accuracy.

A comprehensive survey of privacy preserving and collaborative training for robust steganography was designed in [26]. Yet another method to improve visual imperceptibility employing inception transformer was presented in [27] to not only generate higher quality stego images but also to achieve higher PSNR. A highly secure technique employing Binary bit-plane decomposition (BBPD) based image encryption along with Salp Swarm Optimization Algorithm based embedding process was designed in [28] to ensure better results in terms of security. A method to focus on payload and achieve high PSNR was proposed in [29] employing wavelet transform.

Conventional image steganography mechanisms depend on artificial embedding protocols that as a matter of course manipulate cover images and are prone to steganalysis detection. Recently, reversible data hiding and neural network has made significant progress in steganography. However, existing neural network steganography methods mostly rely on encryption mechanisms that which only possess local receptive fields and constrain data confidentiality. In contrast to previous works, we propose quality enhanced image stegnographic method with spread spectrum called, Functional Segmentation Deep Vision Transformer with Piecewise Spread Spectrum (FSDVT-PSS) for secured communication. This method by using deep vision transformer ensures quality enhancement and on the other hand using spread spectrum makes secure communication. The elaborate description of the FSDVT-PSS method for secure communication is provided in the following sub-section.

3. Methodology

In this work an image stegnographic method called, Functional Segmentation Deep Vision Transformer with Piecewise Spread Spectrum (FSDVT-PSS) for secured communication is proposed. The FSDVT-PSS method uses Deep Vision Transformer (Deep ViT) with Spread Spectrum analysis for secured image transmission. The entire process of FSDVT-PSS method is split into six steps. They are data collection, pre-processing, segmentation, positional encoding, transformer encoder layer and finally classification head.

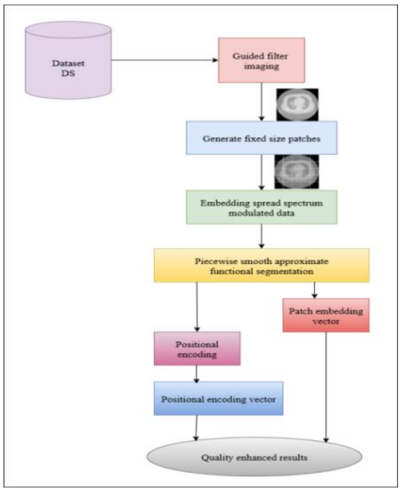


Figure 1: Block diagram of Functional Segmentation Deep Vision Transformer with Piecewise Spread Spectrum (FSDVT-PSS) for secured communication

As shown in the above figure initially, information from Chest CT Scans + 1M DICOM files + reports dataset is employed for performing image steganographic based secure communication ensuring sufficient size and quality for effective training. With the dataset details provided as input is subjected to Guided Filter imaging mechanism to de-noise input medical images.

Next, with the pre-processed samples Patchification along with Segmentation using Piecewise Smooth Approximate Functional Segmentation is applied with the intent of processing each fixed-size patch individually and then combines the results to produce a complete segmentation results. Here, the region with high entropy represents a homogeneous area where segmentation is not required. On the other hand, the region with low entropy represents a region with more complex structure, where segmentation is performed. Accordingly results are stored in Patch Embedding Vector.

Following which positional encoding is performed using sine and cosine encodings, therefore storing the results in Positional Encoding Vector. Next, the Patch Embedding Vector results and Positional Encoding Vector results are then pushed through several Transformer Encoder layers. Finally classification for secure communication is performed using a non-linear activation function. The elaborate

Paper ID: SR251119102846 DOI: https://dx.doi.org/10.21275/SR251119102846

explanation of the image steganography method, FSDVT-PSS for secure communication is provided in the following sub-sections.

3.1 Dataset Selection

Initially in this work, Chest CT Scans + 1M DICOM files + reports dataset [30] extracted from https://www.kaggle.com/datasets/humanaizedata/chest-ct-scans-1m-dicom-files-reports?select=anonym is used, ensuring sufficient size and quality for effective training. The dataset includes over 1 million high-quality DICOM files from thoracic CT scans. To perform image steganographic for secure communication 1418 high-quality DICOM files from a single patient imparting sequential view of thoracic CT scans are employed. Each 1418 high-quality DICOM files from a single patient includes different thorax slice, making certain extensive understandings into the patient's pulmonary anatomy with improved quality enhancement and secure communication.

3.2 Guided Filter Image Pre-processing

Data pre-processing plays a significant role in image steganography for ensuring better quality of data and secure communication between sender and intended recipients. The data or the images being extracted from raw dataset contain noisy data that result in ineffective data analysis and compromises security. In Deep Vision Transformer (ViT), pre-processing task encompasses ascertaining the outliers and accordingly cleaning of noisy data for further processing via the transformer architecture. In this work initially, Guided Filter Image Pre-processing is applied to the raw sample files of Chest CT Scans + 1M DICOM files + reports dataset. The objective is to transform the image and de-noise it with improved peak signal to noise ratio (PSNR). Figure 2 shows the structure of Guided Filter Image Pre-processing.

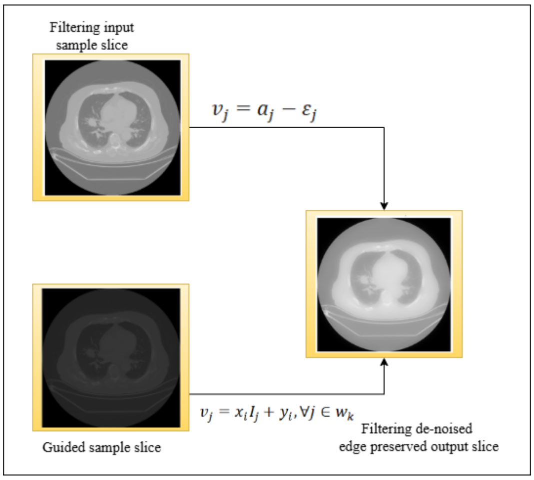


Figure 2: Structure of Guided Filter Image Pre-processing

As shown in the above figure, for each patient 'P' possessing 'M' distinct high-quality DICOM files '$F = \{P_1F_1, P_1F_2, \ldots, P_1F_m\}$', '$m = 1418$' from Dataset '$DS$' the Guided Filter Image Pre-processing model is applied to de-noise input medical images by the transformer architecture. One main hypothesis of the guided filter is that the relation between guidance 'I' and the filtering output 'v' is linear and if 'v' is a linear transformation of in a window 'w_k' centered at the pixel 'k'. In order to determine the linear coefficient '(x_i, y_i)', constraints from the filtering input 'a' are required. Then, the output 'v' is modeled as the input 'a' with undesirable pixels 'n', i.e. noise are subtracted or eliminated by the transformer architecture. Then, the basis guided filter model is formulated as given below..

$$v_j = x_i I_j + y_i, \forall j \in w_k \quad (1)$$
$$v_j = a_j - \varepsilon_j \quad (2)$$

From the above equations (1) and (2) 'v_j' denotes the '$j-th$' output pixel for the corresponding '$j-th$' input pixel 'a_j' with 'ε_j' as the noise component and '$j-th$' guidance pixel 'I_j' respectively. By using this guidance pixel value the edges are preserved extensively. Then, the noise component is evaluated as given below.

$$\varepsilon_j = a_j - x_i I_j - y_i \quad (3)$$

The error function is then formulated as given below.

$$Err(x_i, y_i) = \sum_{j \in w_k} \left(\left(x_i I_j + y_i - a_j \right)^2 + \eta x_i^2 \right) \quad (4)$$

From the above equation (4), 'η' represent the regularization parameter, penalizing input 'x_i' with window 'w' centered at the pixel 'w_k' respectively. Finally, the de-noised input medical images by the transformer architecture are obtained as given below.

$$x_i = \frac{\frac{1}{|w|}\sum_{i \in w_k}\left(I_j a_j - \mu_i a_i'\right)}{\sigma_i^2 + \eta} \quad (5)$$
$$y_i = a_i' - x_i \mu_i \quad (6)$$
$$D_i = (x_i, y_i) \quad (7)$$

With the above equation results from (5) and (6) the de-noised pixels with edge-preserved results 'D_i' by the transformer architecture are obtained for further processing. The pseudo code representation of Guided Filter Image Pre-processing is given below.

Algorithm 1 Guided Filter Image Pre-processing

Input: Dataset 'DS', Sample DICOM CT images or Samples '$S = \{S_1, S_2, \ldots, S_N\}$',
Patients '$P = \{P_1, P_2, \ldots, P_M\}$', '$F = \{P_1F_1, P_1F_2, \ldots, P_1F_m\}$'

Output: De-noised and edge-preserved results

1: **Initialize** 'N', 'M', '$m = 1418$'
2: **Begin**
3: **For** each Dataset 'DS' with Samples 'S', Patients 'P' and patient files 'F'
4: Generate basis guided filter function according to (1) and (2)
5: Formulate noise component according to (3)
6: Generate error function according to (4) for edge preservation
7: Generate de-noised input medical images according to (5) and (6)
8: **Return** de-noised medical images 'D'
9: **End for**
10: **End**

As given in the above algorithm, with the objective of de-noising medical images and also to preserve edges for further processing, initially, basis guided filter function is applied. Following which, to preserve the edge during de-noising process, error function is formulated and accordingly

Paper ID: SR251119102846 DOI: https://dx.doi.org/10.21275/SR251119102846

the results are passed on to the next stage. Finally, by penalizing input image according to regularization parameter, de-noised medical images are obtained with improved peak signal to noise ratio (PSNR).

3.3 Vision Transformer integrated with Spread Spectrum for quality enhanced secure communication

Next, in this section with the de-noised medical images as input, a Vision Transformer model integrated with Spread Spectrum for quality enhanced secure communication is designed. Figure 3 shows the structure of Vision Transformer with Spread Spectrum for quality enhanced secure communication.

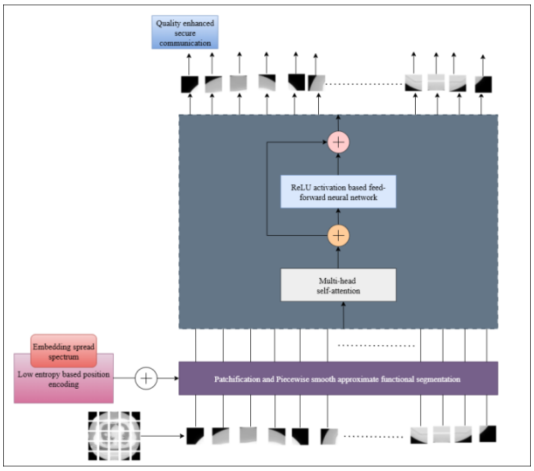


Figure 3: Structure of Vision Transformer with Spread Spectrum for quality enhanced secure communication

As shown in the above figure the de-noised medical images provided as input is subjected to Piecewise Smooth Approximate Functional Segmentation where patchification and segmentation are performed separately and stored as patch embedding vector. Here, areas of image with lower entropy are said to be more suitable for embedding spread spectrum data to ensure secure communication. Following which, low entropy based position encoding stored as positional encoding vector is combined with the patch embedding vector results and pushed through several transformer encoder layers. With each transformer encoder layer consisting of multi-head self-attention are activated via ReLU activation based feed-forward neural network to finally generate the classified results (i.e. enhanced or noisy image) ensuring quality enhancement.

3.3.1 Piecewise Smooth Approximate Functional Segmentation and Spread Spectrum based Patch Embedding

The second step in the Deep Vision Transformers (ViT) is patchfication and segmentation. Here, patchification refers to the process of splitting a de-noised image into smaller, non-overlapping patches and then segmenting them into a sequence of vectors that serve as input tokens for transformer network. In this work, a Piecewise Smooth Approximate Functional Segmentation is applied to the pre-processed de-noised medical for producing Patch Embedding Vector results.

The purpose of using Piecewise Smooth Approximate Functional Segmentation is by splitting de-noised medical images into regions of relatively smooth characteristics and sharp boundaries allow for accurate segmentation of complex images upon comparison to conventional segmentation methods. Each patch here represents local region of de-noise input medical images. Following which Piecewise Smooth Approximate function is applied to process each fixed-size patch individually and then combine the results to produce a complete segmentation of the pre-processed sample image.

Using ViT, pre-processed de-noised medical cover image is analyzed and then optimal locations or frequency bands are ascertained for embedding spread spectrum modulated data. To start with, pre-processed de-noised medical image is split into fixed-size patches for generating patch embedding results. Let us consider as input de-noised medical image with height 'H', width 'W' and channels 'C'. Let us also assume that the patch height 'H' and width 'W' both as 'P' with the de-noised medical image is split into a sequence of '$m = \frac{HW}{P^2}$' patches, where each patch is flattened to a vector of length 'CP^2'.

In this way, image patches generated here can be considered analogous to tokens in text sequences by transformer architecture. Following which Piecewise Smooth Approximate function is applied by approximating de-noised medical image as a combination of smooth regions, therefore handling variations in a more efficient manner. This function, aims at expressing the intensities inside and outside the contour employing the energy function as given below.

$$Energy^{PS}(u^+, u^-, \varphi) = \int |u^+ - D|^2 + H(\varphi)dp + \int |u^- - D|^2 + \left(1 - H(\varphi)\right)dp \qquad (8)$$

From the above equation (8), 'u^+' and 'u^-' denotes the smooth functions approximating the de-noised medical images 'D' inside and outside the contour (i.e. curve) or two regions that are separated by a curve with 'p' representing the image coordinate in the domain of function being segmented.

$$u^+ - D = \mu \Delta u^+ in\, \{p: \varphi(p) > 0\}, \frac{\partial u^+}{\partial n} = 0\ on\, \{p: \varphi(p) = 0\} \qquad (9)$$

$$u^- - D = \mu \Delta u^- in\, \{p: \varphi(p) > 0\}, \frac{\partial u^-}{\partial n} = 0\ on\, \{p: \varphi(p) = 0\} \qquad (10)$$

From the above equations (9) and (10) the smooth functions approximating the de-noised medical images 'D' inside and outside the contour are arrived at using the Poisson function. Finally, the segmented portion is obtained as given below.

$$u(p) = u^+ H(\varphi(p)) + u^+ \left(1 - H(\varphi(p))\right) \qquad (11)$$

For example, a region with high entropy denotes a homogeneous area where the segmentation is less critical. On the other hand, a region with low entropy denotes a region with more complex structure, requiring more careful segmentation. For example, a region with high entropy denotes a homogeneous area where the segmentation is less critical. On the other hand, a region with low entropy denotes a region with heterogeneous area or more complex structure, requiring more careful segmentation. Here, the areas of image with lower entropy or complex patterns are said to be more suitable for embedding spread spectrum signal without causing noticeable artifacts. The entropy in our work is obtained using Renyi entropy as given below.

$$H_\alpha(u) = \frac{1}{1-\alpha}\log\left(\sum_{i=1}^{n} Prob_i^\alpha\right) \quad (12)$$

From the above equation results (12), the areas of image with lower entropy the spread spectrum signal for embedding is performed without causing noticeable artifacts. The results are stored in Patch Embedding Vector.

3.3.2 Low entropy based Positional Encoding Vector

Following which the Positional Encoding is employed to retain spatial arrangement patch information within the image. These positional encodings assists in understanding relative positions of different patches in the segmented medical images. Positional information for understanding spatial relationships within the image is introduced in our work using sine and cosine encodings. This is mathematically represented as given below.

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000^{\frac{2i}{d}}}\right) \quad (13)$$

$$PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{\frac{2i}{d}}}\right) \quad (14)$$

From the above equations (13) and (14) 'pos' refers to the position of the segmented portion with 'd' representing the segmented portion size with which the positional encoding results '$PE_{(pos,2i)}$' and '$PE_{(pos,2i+1)}$' are arrived at. Finally, the patch position is transformed into a vector by positional encoding called Positional Encoding Vector.

3.3.3 Transformer Encoder

Following which the Patch Embedding Vector results and Positional Encoding Vector results are then pushed through several Transformer Encoder layers. Each layer consists of multi-head self-attention for measuring attention weights to prioritize input sequence during encryption process and feed-forward neural network with the intent of capturing global dependencies and refine its understanding. For a given query '$Q \in \mathbb{R}^{Dim_Q}$', a key '$K \in \mathbb{R}^{Dim_K}$' and a value '$V \in \mathbb{R}^{Dim_V}$' each attention head is evaluated as given below.

$$h_i = f\left(WM_i^{(Q)}Q, WM_i^{(K)}K, WM_i^{(V)}V\right) \quad (15)$$

From the above equation (15) '$WM_i^{(Q)}$', '$WM_i^{(K)}$' and '$WM_i^{(V)}$' denote the weight matrix for the corresponding query, key and value with 'f' denoting attention pooling. Then, the multi-head self-attention results are arrived at as given below.

$$mh_i = WM_O\begin{bmatrix} h_1 \\ h_2 \\ \dots \\ h_1 \end{bmatrix} \quad (16)$$

On the basis of the above design as given in (16), each head attend to different parts of the images subsequently. The query here represents the secret data to be hidden whereas key represents the relevancy of secret data to be hidden with respect to the other segmented portions and finally the value represents the actual secret data to be hidden in a particular segmented portion.

Finally, in ViT, feed-forward neural network aids in learning complex relationships between image patches after multi-head self-attention mechanisms process them. These feed-forward neural networks consist of fully connected layers activated via ReLU, a non-linear activation function. By employing the ReLU, a non-linear activation function allows in capturing intricate patterns and representations within the image data where the actual embedding spread spectrum data takes place. This is represented as given below.

$$\sigma(p) = \max(0, p) \quad (17)$$

If we want to classify stego images based on their shape and intensity, by using a line function classification can be done only using shape. However, with the stego image generated is more complex with different form of intensity, by adding ReLU, a non-linear activation function generate curved decision boundaries to classify them correctly. Here, the quality enhanced images obtained using non-linear activation function is then used for secure communication via spread spectrum. The pseudo code representation of Vision Transformer integrated with Spread Spectrum for quality enhanced secure communication is given below.

Table 2 Vision Transformer integrated with Spread Spectrum for quality enhanced secure communication

Input: Dataset 'DS', Sample DICOM CT images or Samples '$S = \{S_1, S_2, \dots, S_N\}$', Patients '$P = \{P_1, P_2, \dots, P_M\}$', '$F = \{P_1F_1, P_1F_2, \dots, P_1F_m\}$'
Output: Quality enhanced secure communication
1: **Initialize** 'N', 'M', '$m = 1418$', de-noised medical images 'D' 2: **Begin** 3: **For** each Dataset 'DS' with Samples 'S', Patients 'P' and patient files 'F' and de-noised medical images 'D' **//Generation of Patch Embedding Vector and obtaining segmentation results** 4: Generate intensity function inside and outside the contour employing energy function according to (8) 5: Formulate smooth functions approximating inside and outside the contour according to (9) and (10) 6: Generate segmented portion according to (11) 7: Generate entropy using Renyi entropy according to (12) 8: Add secret data in low entropy portion 9: Store low entropy enabled secret data results in Patch Embedding Vector **//Generation of Positional Encoding Vector**

10: Generate Positional information using sine and cosine function according to (13) and (14)
11: Store low entropy enabled secret data position in Positional Encoding Vector
//Vision Transformer Encoders
12: Formulate attention head results according to (16)
13: Generate multi-head self-attention results according to (17)
14: **If** '$\sigma(p) \geq 0.5$ and $\sigma(p) < 1$'
15: **Then** quality enhanced image (i.e. stego image)
16: Send quality enhanced image (i.e. stego image) for secure communication
17: **End if**
18: **If** '$\sigma(p) < 0.5$'
19: **Then** noisy image
20: **Go to** step 4
21: **End if**
22: **End for**
23: **End**

As given in the above algorithm, the entire process of Vision Transformer integrated with Spread Spectrum for quality enhanced secure communication is split into four sections. First, with the de-noised medical images provided as input the entire image is split into patches. Following which, Piecewise Smooth Approximate Functional Segmentation is applied approximating inside and outside the contour to generate segmented portions. Next, by employing Renyi entropy function, the areas of image with lower entropy are identified to embed spread spectrum by spreading secret data across wider area rather than employing pseudo random sequence. The low entropy enabled secret data results are stored in Patch Embedding Vector. Second, sine and cosine function are applied to generate positional information. Here, low entropy enabled secret data position is ascertained and stored in Positional Encoding Vector. Third both the Patch Embedding Vector results and Positional Encoding Vector results are pushed through several Transformer Encoder layers. Fourth in Vision Transformer Encoders multi-head self-attention results are generated to send quality enhanced image for secure communication and on the other hand others are discarded from further processing.

4. Qualitative Analysis

In this section case analysis of image steganographic quality enhanced secure communication using robust monument defect detection using Chest CT Scans + 1M DICOM files + reports are simulated by applying Functional Segmentation Deep Vision Transformer with Piecewise Spread Spectrum (FSDVT-PSS) for secured communication. Figure 4 (a) given below provides the sample patient 11 DICOM CT scan input slice for performing simulation.

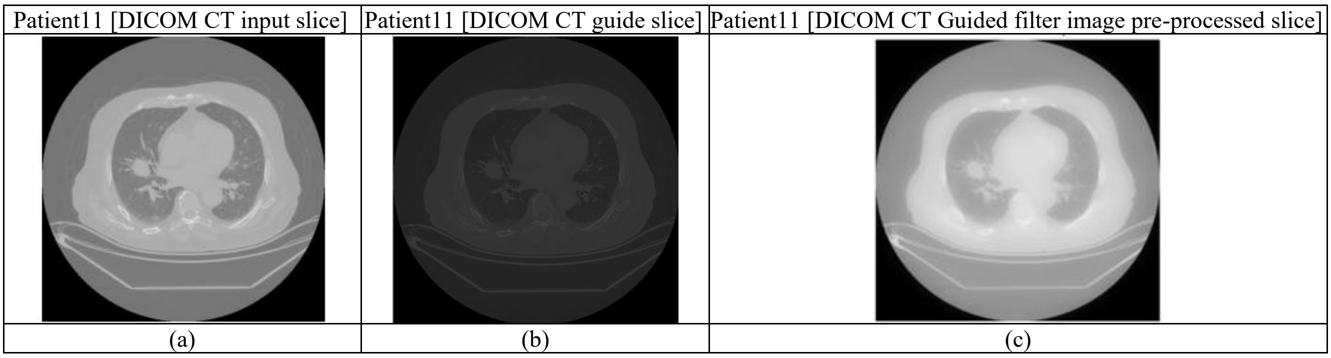


Figure 4 (a) sample image (b) guidance slice image (c) Pre-processed results

As shown in the above figure with the patient 11 DICOM input slice provided as input, Guided Filter Image Pre-processing is applied to generate quality enhanced results. To achieve this, a guidance image as shown in figure 4(b) is provided as input to finally generate de-noised medical images with improved peak signal to noise ratio (PSNR). With the de-noised medical images provided as input are subjected to Vision Transformer with Spread Spectrum for quality enhanced secure communication. Here, initially the de-noised medical images are split into patches as shown in figure 5 (b).

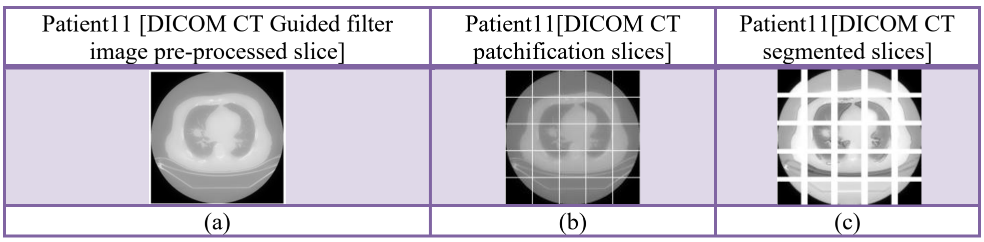


Figure 5: (a) pre-processed results (b) patchified results (c) segmented results

With the above de-noised medical images split into patches (b) are subjected to Piecewise Smooth Approximate Functional Segmentation and Spread Spectrum based Patch Embedding to generate segmented results for further processing. Here by applying the Smooth Approximate Functional Segmentation training time involved in overall classification process is said to be reduced. Finally, figure 6 shows the embedded secret data for secure communication.

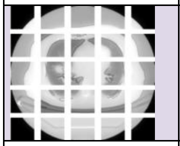
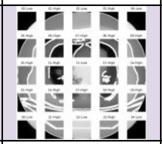
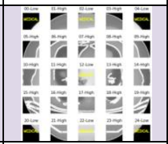
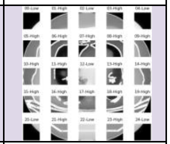
Patient11[DICOM CT patchification and segmented slices]	Patient11[DICOM CT low & high entropy slices using Renyi entropy]	Patient11 - DICOM CT scan low entropy slice with embedded secret data showed slices	Patient11 - DICOM CT scan low entropy slice with embedded secret data hidden slices
			
(a)	(b)	(c)	(d)

Figure 6 (a) segmented slices (b) region with high and low entropy using Renyi function (c) secret data hidden in low entropy (d) encoded data for secure communication

As shown in the above figure, with the segmented slices as input, applying Renyi function high entropy and low entropy results are obtained. Then, secret data is embedded in the low entropy portion. Here, the secret data called, MEDICAL is converted into 8 bits and then the 8-bit data is embedded in the low entropy portion. This in turn aids in improving the data confidentiality and data integrity involved in secure communication process. The quantitative results are discussed in the following sub-sections.

5. Experimental Section and Discussion

This section reports the experimental results for all the three methods, quality enhanced image stegnographic method with spread spectrum called, Functional Segmentation Deep Vision Transformer with Piecewise Spread Spectrum (FSDVT-PSS) for secured communication and existing Reversible Data Hiding method [1], invertible neural network (INN) [2]. The methods are simulated employing Python high-level general-purpose programming on a computer Intel(R) Core (TM) i7-6700HQ CPU@2.60GHz with a RAM of 32 GB running Windows. The dataset used for quality enhanced secure communication is Chest CT Scans + 1M DICOM files + reports extracted from https://www.kaggle.com/datasets/humanaizedata/chest-ct-scans-1m-dicom-files-reports?select=anonym. To ensure fair comparison same dataset is applied to all the three methods. Hence, with the intent of identifying the effectiveness of the method, patient 11 slides are selected arbitrarily. Moreover, several evaluation metrics like, PSNR, training time, data confidentiality, data integrity and accuracy are used to measure the proposed method.

5.1 Case scenario 1: Training time

In this section the training time involved in steganographic quality enhanced secure communication is discussed. The training time is measured as given below.

$$TT = \sum_{i=1}^{N} S_i * Time\ (Class) \quad (18)$$

From the above equation (18) the training time 'TT' is measured based on the samples 'S_i' and the actual time involved in classification process '$Time\ (Class)$'. It is measured in terms of seconds (sec). Table 1 given below lists the training time details using FSDVT-PSS, Reversible Data Hiding method [1] and INN [2]

Table 1 Tabulation for training time

Samples	Training time (sec)		
	FSDVT-PSS	Reversible Data Hiding method [1]	INN [2]
14	0.25	0.28	0.31
28	0.31	0.42	0.53
42	0.35	0.46	0.57
56	0.39	0.5	0.61
70	0.55	0.56	0.67
84	0.68	0.78	0.89
98	0.75	0.86	0.99
112	0.81	0.85	0.96
126	0.85	0.95	1.05
140	0.99	1.05	1.25

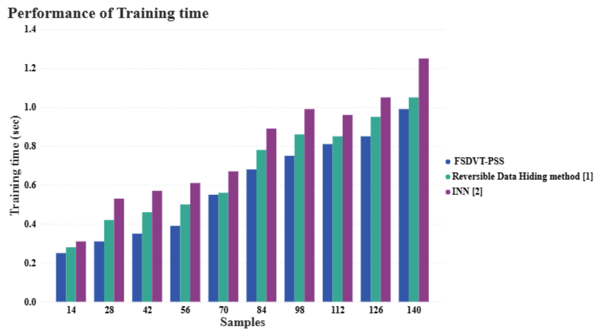


Figure 7: Training time versus samples

Figure 7 given above shows the training time analysis made with 140 different samples using the proposed FSDVT-PSS and two existing methods, Reversible Data Hiding method [1] and INN [2]. Despite increase in training time were observed using all the three methods for the 10 simulation runs, it was found to be comparatively lesser using FSDVT-PSS upon comparison to [1] and [2]. The reason behind the improvement was Piecewise Smooth Approximate Functional Segmentation was applied as patchification and segmentation process by handling intensity with complex backgrounds. This in turn provided a good balance between accuracy and training time. Though CT images employed in our work includes varying intensity within regions (intensity inhomogeneities) can be found challenging for conventional segmentation. In our work by applying the Piecewise Smooth Approximate Functional Segmentation model, by approximating as a combination of smooth regions, can efficiently handle these inhomogeneities. This in turn improves overall training time using proposed FSDVT-PSS method by 16% compared to [1] and 26% compared to [2].

5.2 Case scenario 2: Accuracy

Second in this section the accuracy analysis for secure communication between sender and recipient is analyzed. The accuracy rate is formulated as given below.

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (19)$$

From the above equation (19) accuracy 'Acc' is measured based on the true positive rate 'TP', true negative rate 'TN', false positive rate 'FP' and false negative rate 'FN' respectively. Table 2 given below lists the accurate rate using FSDVT-PSS, Reversible Data Hiding method [1] and INN [2].

Table 2 Tabulation for accuracy

Samples	Accuracy (%)		
	FSDVT-PSS	Reversible Data Hiding method [1]	INN [2]
14	0.96	0.89	0.83
28	0.89	0.82	0.75
42	0.86	0.81	0.74
56	0.84	0.79	0.72
70	0.87	0.82	0.75
84	0.9	0.85	0.78
98	0.92	0.87	0.8
112	0.94	0.89	0.82
126	0.9	0.85	0.78
140	0.95	0.9	0.83

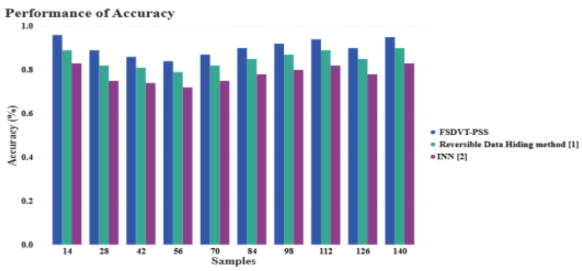


Figure 8 Accuracy versus samples

Figure 8 given above illustrates the accuracy analysis using the three methods, FSDVT-PSS, Reversible Data Hiding method [1], INN [2] respectively. An overall ten simulations were performed using three methods for samples ranging between 14 and 140. From the above graphical representation though the accuracy were found to be decreasing for the first seven iterations whereas for the rest of iterations were found to be improved. Also comparative analysis show improved accuracy rate using FSDVT-PSS method upon comparison to [1] and [2]. The reason was by applying the Piecewise Smooth Approximate Functional Segmentation inhomogeneities were handled efficiently by approximating as a combination of smooth regions. On one hand the smooth functions within each segment encapsulate local intensity differences and on the other hand, the piecewise characteristic permits for distinct intensity in different parts of the image. This in turn improved overall accuracy using FSDVT-PSS method by 6% compared to [1] and 14% compared to [2].

5.3 Case scenario 3: Data confidentiality and Data integrity

Finally in this section the performance analysis of data confidentiality and data integrity rate is evaluated. Data confidentiality makes certain that sensitive information or medical images employed in our work is protected from unauthorized access. On the other hand, data integrity ensures that information is accurate and trustworthy. To be more specific, data (i.e. image) confidentiality protects against unauthorized access, while data (i.e. image) integrity ensures that data accessed is accurate and trustworthy.

$$DC = \sum_{i=1}^{N} \frac{\$_{UA}}{S_i} \quad (20)$$

From the above equation (20) data confidentiality DC rate is measured based on the samples involved in the simulation process "Si" and the samples compromised by unauthorized access "SUC". It is measured in terms of percentage (%). Next, data integrity rate is obtained as given below.

$$DI = \sum_{I=1}^{N} \frac{S_{UC}}{S\,i} \quad (21)$$

From the above equation (21) data integrity "DI" is measured based on the samples involved in the simulation process "Si" and the samples being unchanged "SUC". It is measured in terms of percentage (%). Table 3 given below lists the data confidentiality and data integrity using FSDVT-PSS, Reversible Data Hiding method [1] and INN [2].

Table 3 Tabulation for data confidentiality and data integrity

Samples	Data confidentiality (%)			Data integrity (%)		
	FSDVT-PSS	Reversible Data Hiding method [1]	INN [2]	FSDVT-PSS	Reversible Data Hiding method [1]	INN [2]
14	6.25	8.36	9	4	4.55	4.55
28	7.1	10.71	14.28	3.57	7.1	7.1
42	7.55	11	14.28	3.75	7.25	7.35
56	7.85	11.35	14.35	4	7.55	7.85
70	8.15	11.55	14.55	4.15	7.75	7.95
84	8.35	11.85	14.85	3.55	4	4.25
98	8	10	10.25	3.25	3.55	3.85
112	7.55	9.55	9.75	3	3.35	3.55
126	7.35	9.25	9.55	2.85	3	3.25
140	7	9	0.25	2.55	2.85	3

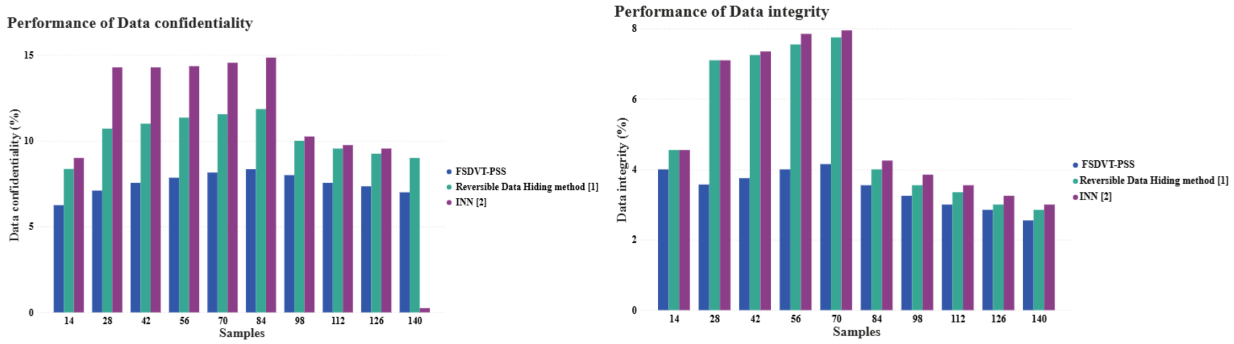


Figure 9: Data confidentiality and data integrity analyses

Figure 9 given above illustrates the data confidentiality and data integrity using three methods, FSDVT-PSS, Reversible Data Hiding method [1] and INN [2]. Here, ten iterations were performed with respect to 140 samples. From the above two figures both data confidentiality and data integrity using FSDVT-PSS method were found to be improved upon comparison to [1] and [2]. In our work, by employing Renyi entropy function, first image portions with lower entropy are identified. The identified area forms the portion where the spread spectrum is applied to spreading secret data across wider area. The low entropy enabled secret data results are stored in Patch Embedding Vector. This results when combined with Position Encoding Vector results aids in improving data confidentiality using FSDVT-PSS method by 36% compared to [1] and 46% compared to [2]. Moreover, sine and cosine function are applied to generate positional information. Also by determining low entropy enabled secret data and storing it in Positional Encoding Vector in turn aids in improving the overall data integrity rate using FSDVT-PSS method by 43% compared to [1] and 48% compared to [2].

6. Conclusion

In this paper, a novel method for quality enhanced image stegnography for secure communication with spread spectrum called, Functional Segmentation Deep Vision Transformer with Piecewise Spread Spectrum (FSDVT-PSS) is proposed. This method targets data confidentiality, data integrity, accuracy and training time as multi objective parameters. The method composed in four parts, Guided Filter Image Pre-processing, Piecewise Smooth Approximate Functional Segmentation and Spread Spectrum based Patch Embedding, Low entropy based Positional Encoding Vector generation, Transformer Encoder with non-linear activation function for efficient classification. First to preserve the edge during de-noising process Guided Filter function was applied that in turn not only improved PSNR but also reduced training time. Next, segmentation process was performed following patch separation using Piecewise Smooth Approximate Function and following which Spread Spectrum based Patch Embedding was done. This in turn improved both data confidentiality and data integrity extensively.

Finally, Low entropy based Positional Encoding Vector generation along with non-linear activation was applied for efficient classification accurately. To validate the efficiency of the proposed method, comparative experimental results are presented by applying Chest CT Scans + 1M DICOM files + reports dataset. The outcomes expressed that the proposed method performed better in terms of the multi objective performance metrics taken into consideration than the current methods Reversible Data Hiding and INN in terms of data confidentiality, data integrity, accuracy and training time.

References

[1] Irsyad Fikriansyah Ramadhan, Ntivuguruzwa Jean De La Croix, Tohari Ahmad and Andre Uzamurengera, “Huffman coding-based data reduction and quadristego logic for secure image steganography”, Engineering Science and Technology, an International Journal, Elsevier, Volume 65, May 2025, Pages 1-15 [Reversible Data Hiding method]

[2] Weida Chen and Weizhe Chen, “Enhanced secure lossless image steganography using invertible neural networks”, Journal of King Saud University - Computer and Information Sciences, Elsevier, Volume 36, Issue 10, December 2024, Pages 1-15 [invertible neural network (INN)]

[3] Saixing Zhou, Miaoxin Ye, Xin Liao, “Color Image Steganography Using Generative Adversarial Networks with a Phased Training Strategy”, ACM Jun 2025

[4] R. Gurunath, Ahmed H. Alahmadi Mohammad Zubair Khan, Debabrata Samanta, Abdul Rahmanal Ahmadi, “A Novel Approach for Linguistic Steganography Evaluation Based on Artificial Neural Networks”, IEEE Access, Vol. 9, Sep 2021

[5] Jiabao Cui, Liangli Zheng, Yunlong Yu, Yining Lin, Huajian Ni, Xin Xu, Zhongfei Zhang, “Deeply-Recursive Attention Network for video steganography”, CAAI Transactions on Intelligence Technology, Wiley, Sep 2022

[6] Zhiyi Wang, Mingcheng Zhou, Boji Liu, Taiyong Li, “Deep Image Steganography Using Transformer and Recursive Permutation”, Entropy, MDPI, Oct 2022

[7] Othman A. Alrusaini, “Deep learning for steganalysis: evaluating model robustness against image transformations”, Frontiers in Artificial Intelligence, Mar 2025

[8] Dr. Ch. Premkumar, A.V.L. Prasuna, Choudari Likhith, P. Sai Akshay, “Steganography: An Enhanced Method for Securely Concealing Information within Digital Image Files”, International Journal of Research Publication and Reviews, Vol. 6, Apr 2025

[9] Yangwen Zhang, Yuling Chen, Hui Dou, Chaoyue Tan, Yun Luo, Haiwei Sang, “Image steganography without embedding by carrier secret information for secure communication in networks”, PLOS ONE, Sep 2024
[10] Heba Ragab, Hassan Shaban, Kareem Ahmed, Abd-Elmgied Ali, “Digital Image Steganography and Reversible Data Hiding: Algorithms, Applications and Recommendations”, Journal of Image and Graphics, Vol. 13, Oct 2025
[11] Samar Kamil Mohammad Kamrulhasan, Siti Norul Huda Sheikh Abdullah, Farah Aqilah Bohani, “Enhanced Flipping Technique to Reduce Variability in Image Steganography”, IEEE Access, Vol. 9, Dec 2021
[12] Hao Li, Yi Zhang, Jinwei Wang, Weiming Zhang, Xiangyang Luo, “Lightweight Steganography Detection Method Based on Multiple Residual Structures and Transformer”, Chinese Journal of Electronics, Vol. 33, Jul 2024
[13] Dhanush Polisetty, Syed Wajahat Abbas Rizvi, “GAN-based Adaptive Image Steganography”, International Journal of Computer Applications, Vol. 187, May 2025
[14] Shahid Rahman, Jamal uddin, Hameed Hussain, Sabir Shah, Abdu Salam, Debora Libertad Ramírez Vargas, Isabel de la Torre Díez, Farhan Amin, Julio César Martínez Espinosa, “A novel and efficient digital image steganography technique using least significant bit substitution”, Scientific Reports, May 2025
[15] Tauqeer Safdar Malik, Kaleem Razzaq Malik, Muhammad Sajid, Ahmad Almogren, Ali Haider Khan, Ayman Altameem, Ateeq Ur Rehman, Seada Hussen, “A hybrid steganography framework using DCT and GAN for secure data communication in the big data era”, Scientific Reports, Jun 2025
[16] Jean De La Croix Ntivuguruzwa, Tohari Ahmad, “A convolutional neural network to detect possible hidden data in spatial domain images”, Cybersecurity, Springer, Nov 2023
[17] Chang Ren, Bin Wu, “A Robust joint coverless image steganography scheme based on two independent modules”, Cybersecurity, Springer, Dec 2024
[18] Dr. M. N. Nachappa, Vignesh Kamble R P, “Image Steganography Applications for Secure Communications”, International Journal of Innovative Science, Engineering & Technology, Vol. 6, May 2019
[19] Fan Li, Chenyang Liu , Zhenbo Dong, Zhibo Sun, Weipeng Qian, “A Robust Coverless Image Steganography Algorithm Based on Image Retrieval with SURF Features”, Security and Communication Networks, Wiley, May 2024
[20] Samira Rezaei, Amir Javadpour, “Bio-Inspired algorithms for secure image steganography: enhancing data security and quality in data transmission”, Multimedia Tools and Applications, Springer, Vol. 83, Mar 2024
[21] Ramyashree, P. S. Venugopala, S. Raghavendra, Vijay S. Kubhihal, “Enhancing Secure Medical Data Communication Through Integration of LSB and DCT for Robust Analysis in Image Steganography”, IEEE Access, Vol. 13, Jan 2025
[22] Eko Heri Susanto, Dimas Pramudya Pratama, Ricki Septian Nurpratama, “Optimizing Digital Image Steganography to Enhance the Security of Secret Message Delivery”, Jan 2024
[23] Zinah Khalid Jasim Jasim, Sefer Kurnaz, “An Improved Image Steganography Security and Capacity Using Ant Colony Algorithm Optimization”, Computers, Materials and Continua, Sep 2024
[24] Oleksandr Kuznetsov, Emanuele Frontoni, Ruslan Shevchuk, Kyrylo Chernov, Mikolaj Karpinski, Kateryna Kuznetsova, “Enhancing Steganography Detection with AI: Fine-Tuning a Deep Residual Network for Spread Spectrum Image Steganography”, Sensors, MDPI, Oct 2024
[25] Osama Alfarraj, Fahad Albelhai, Amr Tolba, Li Bohang, Ningxin Li, Jing Yang, Zaffar Ahmed Shaikh, Roohallah Alizadehsani, Paweł Pławiak, Por Lip Yee, “Image steganalysis using active learning and hyperparameter optimization”, Scientific Reports, May 2025
[26] Fei Yang, Xu Zhang, Shangwei Guo, Daiyuan Chen, Yan Gan, Tao Xiang, Yang Liu, “Robust and privacy-preserving collaborative training: a comprehensive survey”, Artificial Intelligence Review, Springer, Jun 2024
[27] Yunyun Dong, Ping Wei, Ruxin Wang, Bingbing Song, Tingchu Wei, Wei Zhou, “Hiding image with inception transformer”, IET Image Processing, Wiley, Aug 2024
[28] Sachin Dhawan Jaroslav Frnda, Chinmay Chakraborty, Rashmi Gupta, Arun Kumar Rana, Subhendu Kumar Pani, “SSII: Secured and High-Quality Steganography Using Intelligent Hybrid Optimization Algorithms for IoT”, IEEE Access, Vol. 9, Jun 2021
[29] Ravi J , Saleem S Tevaramani, “A Robust Approach to Image Steganography for Secure Communication Using Dual-Tree Complex Wavelet Transformation (DTCWT) and Bit Plane Slicing”, Journal of Systems Engineering and Electronics, Vol. 34, Oct 2024
[30] https://www.kaggle.com/datasets/humanaizedata/chest-ct-scans-1m-dicom-files-reports?select=anonym

Author Profile

V. Arunkumar, PhD Research Scholar (Part Time) in the Department of Computer Science, Periyar University, Salem, Tamil Nadu, India.

Dr. K. Padmanabhan, Controller of Examinations, Vivekanandha College of Arts and Sciences for Women (Autonomous), Tiruchengode, Namakkal, Tamil Nadu, India.