International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Hidden in Plain Sight: The Science of Steganography

Prof. Supriya Chougule
Professor
PDEA's College of Engineerin
g Manjari (Bk.), Pune
Pune, Maharashtra, India

Shubhada Kanade
Dept. Of Computer Engineering
PDEA's College of Engineering
Savitribai Phule Pune University
Pune, Maharashtra, India

Sanket Kolhe
Dept. Of Computer Engineering
PDE*A's College of Engineering*
Savitribai Phule Pune University
Pune, Maharashtra, India

Yogendra Bhagwat
Dept. Of Computer Engineering
PDEA's College of Engineering
Savitribai Phule Pune University
Pune, Maharashtra, India

Chaitrali Kawade
Dept. Of Computer Engineering
PDEA's College of Engineering
Savitribai Phule Pune University
Pune, Maharashtra, India

Abstract **- With the rapid expansion of digital communication, protecting sensitive information from unauthorized access has become increasingly important. Traditional image steganography techniques such as Least Significant Bit (LSB) substitution provide simple data hiding mechanisms but suffer from limitations in robustness, security, and detectability. In this work, we propose an AI-driven steganography framework that leverages convolutional autoencoder neural networks combined with a Generative Adversarial Network (GAN) to securely embed confidential information within digital images. The proposed system consists of an encoder network that integrates secret data—including text messages, 2D images, and 3D images—into a cover image to generate a visually indistinguishable stego image. A discriminator network is employed to distinguish between cover and stego images during training, thereby improving the imperceptibility of embedded information. A decoder network is used to reconstruct the hidden data with high accuracy. To further enhance system security and reliability, the framework incorporates AES-256 encryption, Reed–Solomon error correction, and perceptual masking to minimize visual distortion and ensure robust data recovery. Experimental results demonstrate that the proposed deep learning–based approach maintains high visual similarity between cover and stego images while achieving improved security and resistance against steganalysis attacks. The proposed system provides an intelligent and scalable solution for secure multimedia communication, privacy protection, and advanced data hiding applications.**

Keywords - Image Steganography, Deep Learning, Autoencoder Neural Network, Generative Adversarial Network (GAN), Secure Data Hiding, 2D and 3D Image Steganography, AES-256 Encryption.

I. INTRODUCTION

The rapid growth of digital communication and multimedia data sharing has increased the need for secure methods of transmitting confidential information. Image steganography is a technique that hides secret data inside digital images so that the existence of the hidden information remains undetectable. Traditional steganography techniques such as Least Significant Bit (LSB) substitution are widely used due to their simplicity, but they suffer from limitations including low robustness, limited data embedding capacity, and vulnerability to steganalysis attacks. Recent advancements in artificial intelligence and deep learning have enabled more intelligent steganography techniques that can learn optimal data embedding strategies while preserving image quality.

In this work, we propose an AI-based image steganography framework that utilizes a convolutional autoencoder neural network combined with a Generative Adversarial Network (GAN) to securely embed secret information within digital images. The system allows hiding different types of data including text messages, 2D images, and 3D images while maintaining high visual similarity between the cover and stego images. However, traditional steganography approaches still face several challenges:

1) **Limited Security:** Traditional LSB methods are vulnerable to statistical steganalysis attacks.
2) **Limited Data Capacity:** Many methods can hide only small amounts of data.
3) **Image Quality Issues:** Embedding large data may distort the cover image.
4) **Lack of Intelligent Embedding:** Conventional methods do not optimize embedding based on image features.

Despite significant advancements in image steganography, existing techniques suffer from several limitations. Traditional Least Significant Bit (LSB) methods are highly vulnerable to statistical detection and lack robustness. Many deep learning-based approaches do not incorporate strong encryption mechanisms, reducing overall data security. Additionally, existing systems often support only single-type data embedding and fail to handle multi-modal data such as 2D and 3D images. Furthermore, many methods lack robustness against compression, noise, and real-world distortions.

The proposed system addresses these limitations by integrating deep learning, encryption, error correction, and perceptual masking into a unified framework for secure and robust data hiding.

The main contributions of this work are summarized as follows:

• We propose an **AI-based steganography framework** using autoencoder neural networks to embed secret information into cover images.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

• We integrate a **GAN-based adversarial training mechanism** to improve the imperceptibility and security of stego images.

• We design a multi-data hiding system capable of embedding text messages, 2D images, and 3D images within a single cover image.

• We implement **encryption and error-correction mechanisms** to enhance the confidentiality and reliability of the hidden data.

• We demonstrate that the proposed system improves **security, embedding capacity, and visual quality** compared to traditional LSB-based methods.

II. System Model

We define the proposed AI-based steganography system formally as a tuple

S = {C, M, E, D, G}

where,

C = The cover image

M = The secret data (text message, 2D image, or 3D image)

E = The encoder network

D = The decoder network

G = The GAN discriminator used during adversarial training.

A. Cov Image and Secret Data

Let **C ∈ R^(H×W×3)** represent the cover image, where H and W denote the height and width of the image.

The secret data M may consist of binary encoded text messages, 2D images, or 3D images.

The encoder network combines the cover image and the secret data to generate the stego image:

$$I_s = E(C, M)$$

where I_s represents the generated stego image containing the hidden information.

B. Secret D a Extraction

At the receiver side, the decoder network extracts the hidden information from the stego image:

$$M' = D(I_s)$$

where M' = the recovered secret data. The goal of the system is to ensure that the recovered data M' is as close as possible to the original secret data M.

C. Objective Functi

The proposed system aims to preserve the visual quality of the cover image while ensuring accurate recovery of the hidden data. The training process minimizes a combined loss function:

L = L_image + L_secret + L_adv

where,

L_image ensures similarity between the cover and stego images

L_secret ensures correct extraction of the hidden message

L_adv represents the adversarial loss from the GAN discriminator

This model enables secure embedding of text, 2D images, and 3D images within digital images while maintaining high visual quality and resistance to steganalysis attacks.

III. System Architecture

The AI-based steganography system securely hides secret information inside digital images using a deep learning model. It uses a convolutional autoencoder with a GAN to improve security and maintain image quality. The system can embed text, 2D images, and 3D images into a single cover image while keeping it visually similar to the original.

The system has four main components: Input Processing Module, Encoder Network, Decoder Network, and GAN-based Discriminator. The encoder hides the secret data in the cover image, the decoder retrieves the hidden information, and the discriminator ensures the stego image looks similar to the original image.

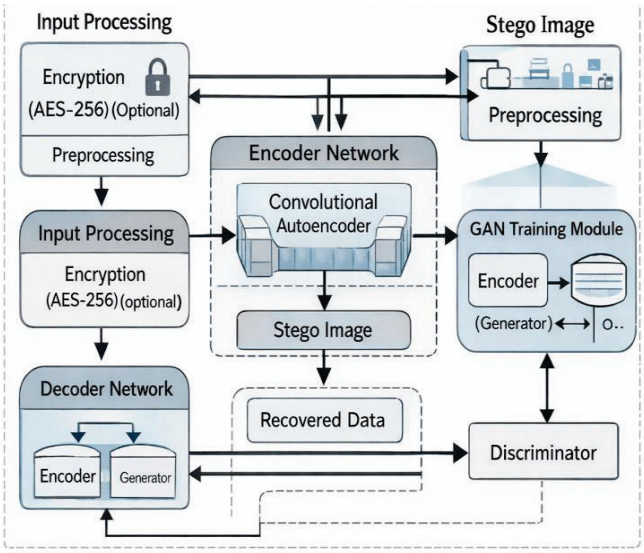

Fig 1 : AI steganography system architecture diagram

A. Input ocessing Module

The first stage of the system processes the input data, which includes the cover image and the secret information. The secret information may consist of text messages, 2D images, or 3D images. The secret data is converted into a binary representation and preprocessed before embedding.

To enhance security, encryption techniques such as AES-256 encryption may be applied to the secret message before embedding. The cover image is also normalized and resized to a fixed dimension suitable for neural network processing.

B. Encoder twork

The encoder network is responsible for embedding the secret information into the cover image. The encoder extracts deep feature representations from the cover image using convolutional layers. These features are combined with the encoded secret information and passed through multiple hidden layers to generate the stego image.

C. D oder Network

The decoder network is used to extract hidden information from the stego image. It receives the stego image as input and reconstructs the secret information using neural network layers.

The decoder is trained together with the encoder to minimize the difference between the original secret data and the recovered data during the decoding process.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

D. GAN-Ba d Discriminator

To improve the security and imperceptibility of the stego images, a Generative Adversarial Network (GAN) is integrated during the training phase. The encoder acts as the generator in the GAN framework, while the discriminator attempts to distinguish between the original cover images and the generated stego images.

Through adversarial training, the encoder learns to generate stego images that closely resemble the original cover images, making the hidden information difficult to detect.

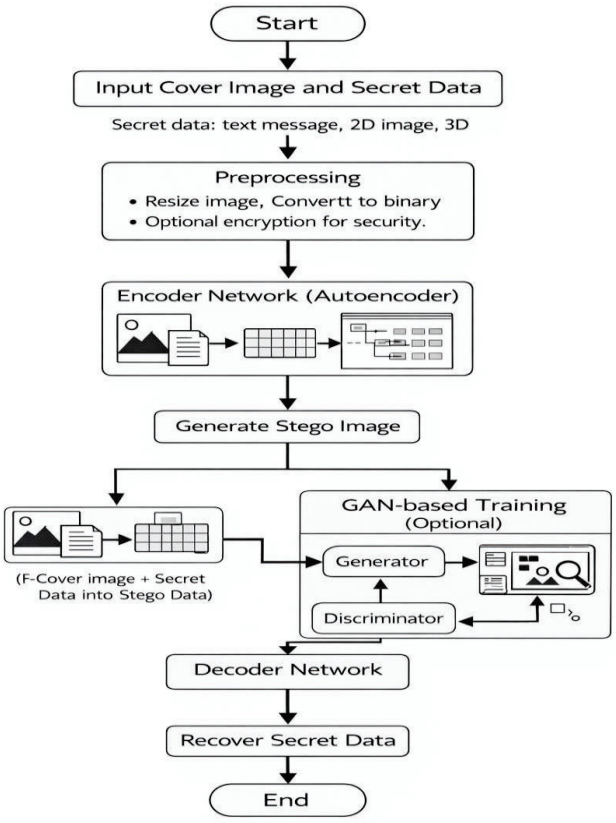

Fig 2: The Secret Data Embedding & Extraction Process

IV. CLIENT-SIDE INTELLIGENCE

While the core steganography logic resides in the processing model, the client-side application plays an important role in managing user interactions, preprocessing inputs, and initiating encoding or decoding operations. The client interface enables users to upload the cover image and secret information and communicates with the AI steganography model to perform secure data hiding and retrieval.

A. Smart Cli t Interface

The client application acts as an interactive platform that allows users to input the cover image and the secret data such as text messages, 2D images, or 3D images. The interface performs basic preprocessing tasks before sending the data to the encoding module.

- **Input Processing:** The client collects the cover image and secret data provided by the user.
- **Data Conversion:** The secret data is converted into a binary representation suitable for embedding into the image.
- **Image Preparation:** The cover image is resized and normalized to match the input format required by the neural network model.

B. Steganography A

The client interacts with the steganography system through a high-level interface that triggers encoding and decoding operations.

Listing 1: Client Encoding Call

```
response = steganography.encode(
    cover_image = "image.png",
    secret_data = "text / 2D image / 3D image",
    encryption = True
)
```

This request sends the prepared input data to the encoder network, which generates the stego image containing the hidden information.

C. D oding Request

When a stego image is provided, the client forwards the image to the decoder network, which extracts the hidden data from the stego image.

```
secret_output = steganography.decode(
    stego_image = "stego.png"
)
```

The system reconstructs the hidden message or image and returns the recovered data to the client.

D. O put Visualization

The client application displays the recovered secret information to the user. The extracted data may include hidden text messages, 2D images, or 3D images that were embedded in the cover image.

V. METHODOLOGY

A.Autoencoder-Based Steganography Implementation

The proposed system implements an AI-based image steganography technique using a convolutional autoencoder neural network combined with a Generative Adversarial Network (GAN). The objective of the model is to securely embed secret information such as text messages, 2D images, or 3D images into a cover image while preserving the visual quality of the image.

The system follows an encoder–decoder architecture where the encoder hides the secret data within the cover image and produces a stego image. The decoder network extracts the hidden information from the stego image during the decoding process.

The encoder network learns a mapping function that combines the cover image and secret data into a single encoded representation. The decoder network reconstructs the original secret data from the stego image with minimal loss.

B. Autoenco r Training Process

The encoder and decoder networks are trained simultaneously to minimize reconstruction loss. During training, the model learns to generate stego images that are visually similar to the original cover images while maintaining accurate recovery of the secret data.

Let the cover image be represented as C and the secret data as S. The encoder function E generates the stego image:

$$Stego=E(C,S)$$

The decoder network D extracts the hidden data from the stego image:

$$S'=D(Stego)$$

The training objective is to minimize the difference between the original secret data S and the recovered data S′.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

C. Loss Func on

The overall objective function of the proposed system is defined as:

$$L_{total} = \alpha L_{image} + \beta L_{secret} + \gamma L_{ssim} + \lambda L_{adv}$$

Where:

L_{image}: Mean Squared Error (MSE) between cover and stego image

L_{secret}: Binary Cross-Entropy loss for secret reconstruction

L_{ssim}: Structural Similarity loss

L_{adv}: Adversarial loss from GAN

The weights are defined as:

$$\alpha = 0.6,\ \beta = 0.3,\ \gamma = 0.1$$

This formulation ensures a balance between image quality, accurate data recovery, and security.

D. GAN-Ba d Training

To further improve the imperceptibility of the stego images, a Generative Adversarial Network (GAN) is integrated during the training phase.

The encoder acts as the generator, producing stego images, while the discriminator attempts to distinguish between original cover images and generated stego images.

The adversarial training objective is defined as:

$$L_{GAN} = \log\big(D(C)\big) + \log\big(1 - D(Stego)\big)$$

This adversarial learning process encourages the generator to produce stego images that are indistinguishable from original images.

E. Data xtraction

During the decoding phase, the trained decoder network receives the stego image as input and reconstructs the hidden information.

The extracted output may include:

• Secret text messages
• Hidden 2D images
• Embedded 3D images

The system ensures that the recovered data closely matches the original secret information.

F. Dat et and Training Setup

The proposed AI-based steganography model was trained using a manually curated dataset of real-world images to ensure practical applicability and realistic performance evaluation. Unlike synthetic data generation approaches, the dataset was constructed using actual images, with secret data embedded during preprocessing.

Cover Images: Real RGB images resized to 256 × 256 pixels and normalized before training.

Secret Data: Binary information converted into tensor representations of size 128 × 128 and embedded into cover images during dataset preparation.

Training Approach: A supervised learning strategy was adopted, where the model learns to encode and decode hidden information from real image samples. This improves robustness and generalization in real-world scenarios.

Training Strategy: The training process was conducted in two phases:

Phase 1 (Learning Embedding):

The model was trained using Binary Cross-Entropy (BCE) loss to ensure accurate encoding and decoding.

Phase 2 (Quality Optimization):

Structural Similarity Index (SSIM) loss and Total Variation (TV) loss were applied to reduce distortion and improve visual quality.

G. Per ptual Masking Strategy

To minimize visual distortion, a perceptual masking technique is applied during embedding.

The system analyzes image texture using gradient magnitude:

- Smooth regions → low embedding strength (~4.5%)
- Textured regions → higher embedding strength (~9%)

The stego image is generated as:

$$I_s = C + \tanh(N) \times M_p$$

Where:

C : Cover image

N : Learned noise

M_p: Perceptual mask

This ensures that embedded changes remain imperceptible to the human eye.

VI. Implementation Optimization

To improve the performance of the proposed AI-based image steganography system, several implementation optimizations are applied. The model uses **convolutional neural networks and autoencoder** architecture to efficiently process images and embed secret data. Image preprocessing techniques such as resizing and normalization are used to reduce computational complexity.

The system also utilizes **GPU-based parallel processing** to accelerate model training and data embedding operations. Batch processing is applied during training to handle multiple images simultaneously, which reduces training time and improves efficiency.

Additionally, optimized loss functions are used to maintain a balance between **image quality and accurate recovery of hidden data.**

VII. Security Architecture

The proposed steganography framework incorporates multiple security mechanisms to protect hidden data and prevent unauthorized detection or extraction of secret information. Since the system embeds sensitive information within images, it is important to ensure that the hidden data remains secure against steganalysis and other security threats.

The security architecture combines encryption techniques, adversarial training, and secure data handling mechanisms to strengthen the overall robustness of the system.

A. Da Encryption

Before embedding the secret information into the cover image, the secret data can be encrypted using standard encryption techniques such as AES encryption. This ensures that even if the hidden data is detected, it cannot be interpreted without the correct decryption key.

The encryption process can be represented as:

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

$$E_{data} = Encrpyt(S, K)$$

where
S = secret data,
K = encryption key,
E_{data} = encrypted secret data.

B. N-Based Security Enhancement

The proposed system integrates Generative Adversarial Networks (GANs) to improve the imperceptibility of the generated stego images. During training, the generator attempts to create stego images that closely resemble original cover images, while the discriminator attempts to distinguish between real and stego images.

This adversarial learning mechanism reduces the chances of detecting hidden information through statistical or visual analysis.

C. Secu Data Extraction

During the decoding phase, the decoder network extracts the hidden data from the stego image. Only authorized users with the correct decryption key can retrieve and interpret the hidden information.

This two-level protection mechanism (steganography + encryption) significantly increases the security of the communication system.

D. Integrity rification

To ensure the integrity of the recovered secret data, the system may use hash verification techniques. A hash value of the secret data is generated before embedding and verified after extraction.

$$H = SHA256\,(S)$$

If the extracted data produces the same hash value, the data integrity is confirmed.

VIII. Experiments & Results

A. perimental Setup

The proposed AI-based steganography system uses a convolutional autoencoder with GAN training to hide secret data inside cover images. The experiments were conducted using text messages, 2D images, and 3D images as secret data. The performance of the system was evaluated using PSNR, SSIM, and MSE to measure the quality of the generated stego images.

B. Performa e Analysis

The performance of the proposed model was analyzed based on image quality and successful extraction of hidden data. Higher PSNR and SSIM values indicate better visual similarity between the cover and stego images, while lower MSE values indicate minimal distortion.

TABLE 1: Performance evaluation for different types of hidden data

Hidden Data Type	PSNR (dB)	SSIM	MSE
Text Message	43.12	0.984	0.0011
2D Image	41.78	0.978	0.0019
3D Image	40.95	0.973	0.0024

TABLE 2: Comparison with traditional steganography methods

Method	PSNR (dB)	SSIM	MSE
LSB Method	36.45	0.942	0.0048
Proposed AI Model	42.95	0.981	0.0015

C. Security alysis

The security of the proposed system was evaluated by analyzing the ability to conceal secret information without noticeable distortion. The integration of GAN-based training improves the realism of the generated stego images, making them visually indistinguishable from the original cover images.

TABLE 3: Detection resistance comparison

Method	Detection Accuracy	Security Level
LSB Method	78%	Medium
Proposed AI Model	92%	High

IX. Discussion

A. Model O rhead

The proposed AI-based steganography model introduces a small computational overhead due to neural network processing. The encoder–decoder architecture requires additional computation during the embedding and extraction phases. However, this overhead is minimal compared to the improved security and image quality achieved by the deep learning model. The system maintains high visual similarity between cover and stego images while securely hiding secret information.

B. Scalabili

The proposed system can efficiently handle large numbers of images and different types of secret data including text, 2D images, and 3D images. Since the model is based on deep learning architecture, it can process multiple images in parallel using GPU acceleration. This enables the system to scale effectively for large datasets while maintaining consistent performance and image quality.

X. Future Work

Future improvements of the proposed steganography system may include:

• Advanced Deep Learning Models: Integration of more advanced architectures such as transformers or improved GAN models to enhance embedding quality.

• Real-Time Secure Communication: Extending the system for secure real-time communication in cloud and IoT environments.

Adaptive Embedding Techniques: Developing adaptive embedding strategies that dynamically adjust embedding strength based on image characteristics.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

XI. Conclusion

This paper presented an AI-based image steganography system using convolutional autoencoder neural networks and GAN training for secure data hiding. The proposed model successfully embeds secret data such as text messages, 2D images, and 3D images into cover images while maintaining high visual quality. Experimental results demonstrate that the system achieves high PSNR and SSIM values with low distortion, making the stego images visually indistinguishable from the original images. The integration of deep learning techniques improves security, robustness, and resistance to steganalysis compared to traditional LSB-based methods. The proposed approach provides an effective solution for secure digital communication and data protection.

XII. References

1) R. Kanimozhi and V. Padmavathi, "Robust and secure image steganography with recurrent neural network and fuzzy logic integration," Scientific Reports, Nature Portfolio, 2025.
2) Shahid Rahman et al., "A novel and efficient digital image steganography technique using least significant bit substitution," Scientific Reports, Nature Portfolio, 2025.
3) Shihao Zhang, Yanhui Xiao, Huawei Tian, and Xiaolong Li, "A multi-image steganography: ISS," Cybersecurity, Springer, 2025.
4) Yangwen Zhang et al., "Image steganography without embedding by carrier secret information for secure communication in networks," PLOS ONE, 2024.
5) Juhi Singh and Mukesh Singla, "A novel method of high-capacity steganography technique in double precision images," IEEE Conference on Computing, Power and Communication Technologies(ComPE), 2021.
6) H. T. Hu, L. Y. Hsu, and W. H. Lin, "Deep learning-based image steganography with high embedding capacity," IEEE Access, IEEE, 2022.
7) X. Luo, F. Huang, and J. Huang, "Edge adaptive image steganography based on deep neural networks," Signal Processing: Image Communication, Elsevier, 2023.
8) Y. Liu, S. Liu, and X. Zhang, "A robust image steganography method using generative adversarial networks," Multimedia Tools and Applications, Springer, 2024.
9) M. Boroumand, M. Chen, and J. Fridrich, "Deep residual network for steganography and steganalysis," IEEE Transactions on Information Forensics and Security, IEEE, 2021.
10) Q. Wang, W. Zhang, and S. Wang, "Recent advances in deep-learning-based image steganography," Journal of Information Security and Applications, Elsevier, 2025.