International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

To compare and analyze the effectiveness of Various state of the art Stegno GAN Methods

A.Dharani[1], M.Lathika Sri[2], T.Muneeswari[3], J.Hemalatha[4], K.Anuradha[5], M.Sekar[6]

[1,2,3]Students, Department of CSE, AAA College of Engineering and Technology, Amathur, Sivakasi, TamilNadu, India
[4,5]Professor, Department of CSE, AAA College of Engineering and Technology, Amathur, Sivakasi, TamilNadu, India
[6]Professor, Department of Mechanical Engineering, AAA College of Engineering and Technology, Amathur, Sivakasi, TamilNadu, India

Abstract

Image steganography is a procedure for hiding messages inside pictures. While other techniques such as cryptography aim to prevent adversaries from reading the secret message, steganography aims to hide the presence of the message itself.With the increasing demand for secure data transmission, deep learning-based data hiding techniques have emerged as promising solutions. This paper explores the Effectiveness of Generative Adversarial Networks (GANs) for data hiding, focusing on their ability to embed secret information into images while maintaining imperceptibility and robustness. Our proposed method utilizes a GAN architecture where the generator learns to embed data seamlessly into images, while the discriminator ensures that the stegno images remain indistinguishable from real images. Experimental results demonstrate the effectiveness of our approach in achieving high imperceptibility and resistance to common image processing attacks.

Keywords: Data hiding, Steganography, Generative Adversarial Networks, Deep Learning, Information Security

1. Introduction

The word steganography is derived from the Greek words stegos meaning cover and grafia meaning writing defining it as covered writing. Image steganography the information is hidden exclusively in images. Steganography is the art and science of secret communication. It is the practice of encoding/embedding secret information in a manner such that the existence of the information is invisible. The actual files can be referred to as cover text, the cover image, or cover audio message. After inserting the secret message it is referred to as stego medium. A stego-key has been used for hiding encoding process to restrict detection or extraction of the embedded data.

Data hiding is a crucial technique in information security, allowing the embedding of secret data within digital media without significantly altering its perceptual quality. Traditional methods such as Least Significant Bit (LSB) steganography and Discrete Cosine Transform (DCT)-based approaches have limitations in robustness and security. Recent advancements in deep learning, particularly Generative Adversarial Networks (GANs), have opened new avenues for enhancing data hiding techniques. This paper presents a GAN-based approach for embedding secret messages into images while preserving visual quality and ensuring robustness against attacks.

The internet revolution offers ease in digital communication; at the same time, it is also a challenge for us to secure the message over the open network. The security system plays a vital role to restrict the messages from being seized by an unauthorized person. Cryptography protects the content of the information that allows only the sender and intended beneficiary of communication to view its contents. They are used to obscure the confidential information within the innocent media like image, video, audio, and text, information. On the basis of the capability of recovering the cover images, data hiding techniques are categorized into two groups: irreversible and reversible data hiding (RDH). If the cover image can be obtained after removal of the confidential data the process is said to be reversible data hiding; otherwise it is termed as irreversible data hiding.

Traditional approaches to image steganography are only effective up to a relative payload of around 0.4 bits per pixel (Pevny et al. ´ , 2010). Beyond that point, they tend to introduce artifacts that can be easily detected by auto mated steganalysis tools and, in extreme cases, by the hu man eye. With the advent of deep learning in the past decade, a new class of image steganography approaches is emerging (Hayes & Danezis, 2017; Baluja, 2017; Zhu et al., 2018). These approaches use neural networks as either a compo nent in a traditional algorithm.identify spatial locations suitable for embedding data), or as an end-to-end solution, which takes in a cover image and a secret message and combines them into a steganographic image.

These attempts have proved that deep learning can be used for practical end-to-end image steganography, and have achieved embedding rates competitive with those accom plished through traditional techniques (Pevny et al. ´,2010). However, they are also more limited than their traditional counterparts: they often impose special constraints on the size of the cover image (for example, (Hayes & Danezis, 2017) requires the cover images to be 32 x 32); they attempt to embed images inside images and not arbitrary messages or bit vectors; and finally, they do not explore the limits of how much information can be hidden successfully.

DE, histogram shifting, and Interpolation based techniques, etc., are the examples of reversible data hiding techniques whereas LSB, PVD, etc., are examples of irreversible data hiding techniques. Conversely, steganalysis, is the art of identifying the hidden information embedded in digital media. After the heart-breaking incidents of September 2001, researchers have given great importance to the topic steganography and steganalysis. It has also now become an important research topic due to the popularity of social media applications like Facebook, WhatsApp, etc.

2. Abbreviations and Acronyms

- GAN - Generative Adversarial Network.
- LSB - Least Significant Bit.
- DCT - Discrete Cosine Transform.
- RDH - Reversible Data Hiding.
- DCNN - Deep Convolutional Neural Network.
- PIL - Python Imaging Library.
- Conv2D - 2D Convolutional Layer.
- ReLU - Rectified Linear Unit.
- PSNR - Peak Signal-to-Noise Ratio.
- SSIM - Structural Similarity Index Measure.

3. Problem Statement

With the rising need for secure communication, existing steganographic methods face significant challenges in maintaining a balance between imperceptibility, payload capacity, and security. Many traditional techniques introduce noticeable distortions in cover images or fail to withstand sophisticated steganalysis attacks, limiting their effectiveness. Additionally, the adaptability of conventional methods to diverse datasets and real-world scenarios remains inadequate. In modern steganography, maximizing secret message capacity while ensuring minimal perceptual distortion and maximum security remains a challenge.

4. Objectives

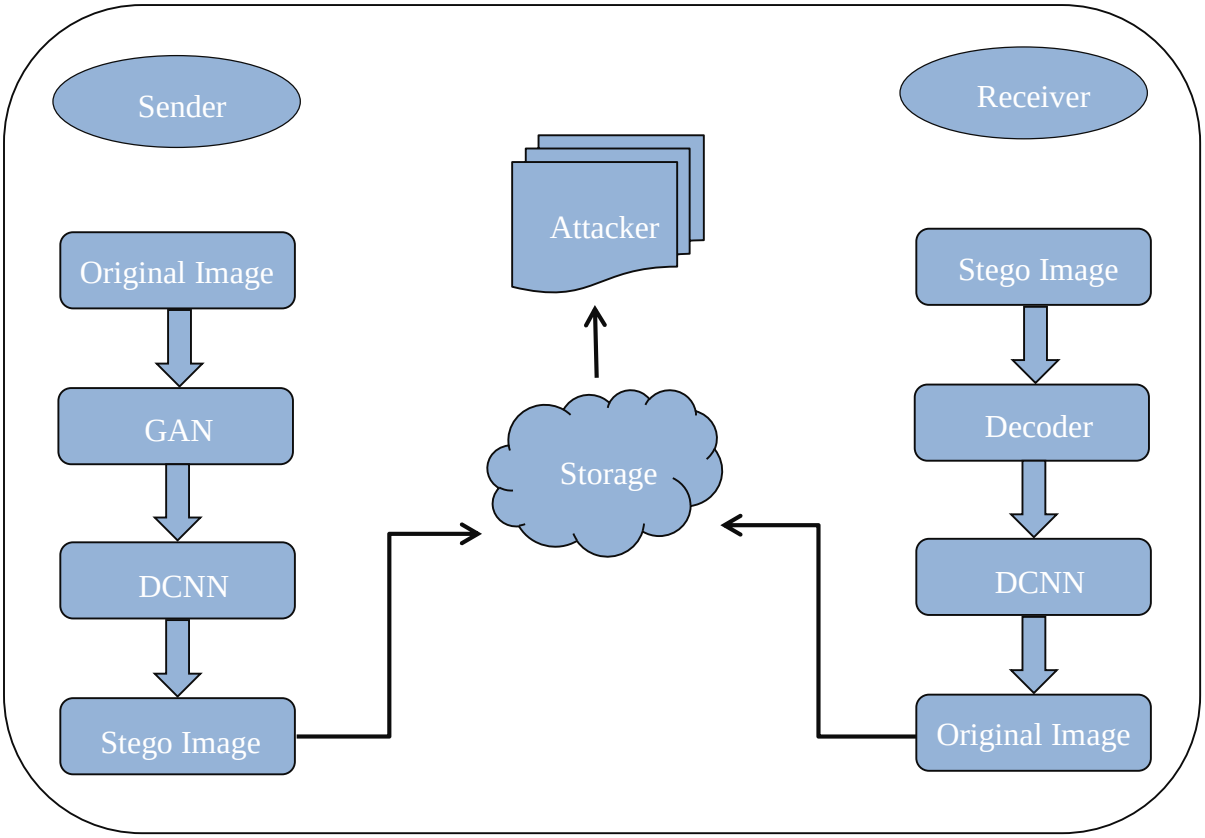
This research aims to develop an advanced GAN-based steganography model that enhances security, maximizes embedding capacity, and ensures high-quality image reconstruction while remaining undetectable by modern steganalysis techniques. By leveraging deep learning and generative models, the proposed approach offers a more resilient and efficient data-hiding technique, setting a new benchmark in secure digital communication.

5. Scope

The training process involves:

1. Feeding a cover image and secret message into the generator.
2. The generator producing a stego image that closely resembles natural images.
3. The discriminator evaluating the authenticity of the generated image.
4. The decoder extracting the hidden message while minimizing reconstruction loss.
5. Optimization using adversarial loss, perceptual loss, and reconstruction loss.

6. Diagram for GAN Architecture



7. Explanation

- Sender:

The sender is the party that embeds a secret message into a cover image using a neural network (generator). It takes both the message and the cover image as inputs and outputs a stego image that looks natural to human eyes but contains hidden information.

- Storage:

This represents the medium through which the stego image is stored or transmitted. It could be cloud storage, email, social media, or any digital channel. The goal is to ensure the stego image remains visually unchanged and undetected during storage or transmission.

- Reciver:

The receiver is the party who receives the stego image and uses a trained decoder network to extract the hidden message from it. The success of this step depends on how well the message was embedded and how robust the model is to distortion or noise during transmission.

- DCNN:

This refers to the core deep learning components including the generator and discriminator.

- The generator learns to embed the message into an image.
- The discriminator tries to distinguish between natural and stego images, improving the realism of outputs.

- Decoder:

The decoder is a neural network trained to extract the hidden message from the stego image. It aims to recover the message accurately by minimizing reconstruction loss, even if the stego image undergoes slight distortions during transmission.

8. Methodology

- GAN Architecture for Data Hiding

Our proposed framework consists of three primary components:

1. Generator: A deep convolutional neural network (DCNN) that learns to embed secret data into images.
2. Discriminator: A network that distinguishes between real and stego images, ensuring high imperceptibility.
3. Decoder: Extracts the embedded data from stego images with minimal errors.

We train our model using an adversarial loss function, a perceptual loss to preserve image quality, and a reconstruction loss to enhance data extraction accuracy

- **Training Process**

The training process involves:

1. Feeding a cover image and secret message into the generator.
2. The generator producing a stego image that closely resembles natural images.
3. The discriminator evaluating the authenticity of the generated image.

4. The decoder extracting the hidden message while minimizing reconstruction loss.
5. Optimization using adversarial loss, perceptual loss, and reconstruction loss

9. Algorithms

1. Import necessary libraries:

- Tensorflow for training the Generative Adversarial Network (GAN).
- Numpy for handling arrays and image data manipulation.
- Pillow (PIL) for Python Imaging Library is used to opening, manipulating, and saving image files.

2. Input a cover image (c) and a secret message (w).
3. Convert the text (m) into an image format (text-to-image).
4. Resize both cover image and message image to a fixed size (e.g,256x256).
5. Normalize pixel values of both images to range [0, 1].
6. Concatenate the cover image and the message image.
7. Feed the concatenated image into a DCNN-based Generator:

- Use layers: Conv2D, ReLU, BatchNormalization, etc.

8. Generator outputs a stego image (s) that looks like the original cover image.
9. Pass the stego image and real images through a Discriminator (DCNN).
10. Discriminator tries to distinguish between real and generated (stego) images.
11. Discriminator is used to train the generator adversarially.
12. A separate Decoder is used to extract the hidden message (m') from the stego image (s).
13. The Decoder (DCNN) is used to extract the hidden message (m') from the stego image (s).
14. Use the following loss functions:

- Adversarial Loss: For realism of stego image.
- Reconstruction Loss: For accurate message extraction.
- Perceptual Loss (optional): For high-level visual similarity.

15. Train the Generator, Discriminator, and Decoder iteratively using backpropagation and Adam optimizer.
16. Save the final stego image (S) (e.g., generated_stego.png).
17. To extract the message later, load (s), feed into the decoder, and reconstruct (m).
18. End.

10. Visual Experimental Result

To assess the impact of the proposed method,three visual outputs were generated:

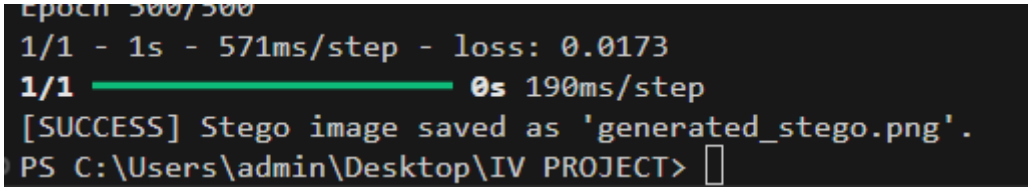
1.Original Image: Input image before message embedding.

2.Stego Image: The image after embedding the encrypted message.

3.Decoded Image: The final output after applying GAN Technology.

11. Results

The proposed model successfully generated a stego image that visually resembles the original cover image with minimal perceptual distortion. The secret message was effectively embedded using a GAN-based architecture, ensuring imperceptibility and high fidelity.



References

1. AGASI: "A Generative Adversarial Network-Based Approach to Strengthening Adversarial Image Steganography" is Authored by Haiju Fan, Changyuan Jin, Ming Li (2025) This paper introduces AGASI, a GAN-based method designed to enhance the robustness of stego-images against steganalysis tools.
2. "Deep Convolutional Generative Adversarial Network for Image Steganography Enhancement" is Authored by Yugandhar Gara, S. I. F. (2024). This paper presents a DCGAN-based approach to enhance image steganography, focusing on improving the quality and security of stego-images. The method leverages deep convolutional networks to achieve better embedding performance.
3. "VidaGAN: Adaptive GAN for Image Steganography" is Authored by Ramandi et al. (2024) VidaGAN introduces an adaptive GAN model for image steganography, balancing embedding capacity and image quality. The approach allows for adjustable payload sizes while maintaining high recovery accuracy and visual fidelity.
4. "GAN-based Image Steganography for Enhancing Security via Adversarial Attack and Pixel-wise Deep Fusion" Authors*:* Chao Yuan, Hongxia Wang, Peisong He, Jie Luo, Bin Li Published on (2022) in Multimedia Tools and Applications. This paper proposes an end-to-end image steganographic scheme based on GANs, incorporating adversarial attacks and pixel-wise deep fusion to enhance security against CNN-based steganalyzers.
5. "SteganoCNN: Image Steganography with Generalization Ability Based on Convolutional Neural Network" is Authored by Xintao Duan et al. (2020) SteganoCNN presents a CNN-based steganography method emphasizing generalization capabilities. The model is trained on diverse datasets to ensure robustness across various image types, enhancing its applicability in real-world scenarios.
6. "Color Image Steganography Using Deep Convolutional Autoencoders Based on ResNet Architecture" is Authored by Seyed Hesam Odin Hashemi et al. (2022). The authors introduce a steganography scheme combining deep convolutional autoencoders with ResNet architecture. This approach aims to improve the capacity and imperceptibility of hidden images, achieving high PSNR and SSIM values.
7. "Generative Steganography Diffusion" is Authored by Authors: Ping Wei, Qing Zhou, Zichi Wang, Published on May (2023) This paper presents a generative steganography method using diffusion models, achieving high-quality stego images and exact recovery of hidden data through an invertible diffusion process.