Generative Image Steganography

An information hiding tool using "Generative Image Steganography" methods to avoid detection by steganalizers.

Erdem Yağcı, Yunus Emre Keleş Ahmetcan Oğuz,
Supervisor: Assoc. Prof. Dr. Harun Artuner
Hacettepe University

Introduction

- This project focuses on developing an innovative information-hiding tool utilizing "Generative Image Steganography" techniques to securely embed data within digital images. By leveraging generative models such as GANs (Generative Adversarial Networks), the tool will create highly imperceptible and realistic images that conceal sensitive information. Unlike traditional steganography methods, generative approaches aim to produce cover images that seamlessly integrate embedded data, making detection by modern steganalysis systems significantly more challenging.
- The project will encompass the design and implementation of a system capable of:
  1. Encoding secret data into images while preserving visual quality as much as possible.
  2. Decoding the embedded data with high accuracy possible.
- The project's results have potential applications in secure communication, digital watermarking, and protecting intellectual property. Through this work, we aim to advance the field of information hiding by exploring cutting-edge generative methodologies to enhance privacy and data security.

Project Description

Steganography is the practice of concealing information within digital media, such as images, to ensure secure communication. Traditional methods often embed secret data into existing cover images by modifying their pixel values or frequency components. However, these methods are increasingly vulnerable to detection by modern steganalysis systems, which employ machine learning techniques to identify subtle anomalies introduced by embedding processes.

The challenge lies in developing a method that produces highly imperceptible steganographic images, resistant to detection by both statistical and learning-based steganalyzers. Generative Image Steganography offers a promising solution by leveraging generative models, such as Generative Adversarial Networks (GANs), to create synthetic images that naturally embed secret data without introducing detectable artifacts.

The rise of adversarial technologies, such as advanced steganalysis tools, necessitates innovative steganography techniques. Existing studies demonstrate the potential of GAN-based methods to outperform traditional approaches in imperceptibility and robustness. Generative Image Steganography creates a new paradigm by directly generating carrier images, optimizing them for both data embedding and security.

Expected Outputs

1. Stego Images: Realistic images containing hidden data, indistinguishable from natural images to both human observers and automated detection systems.

2. Extracted Data: The secret data retrieved with minimal or no loss.

Conclusions: Our current implementation embeds and decodes 100-bit vector data. The capacity of this model can be increased. We also get an accuracy of around 85 percent. The model can be updated and retrained to higher levels.

Resources:

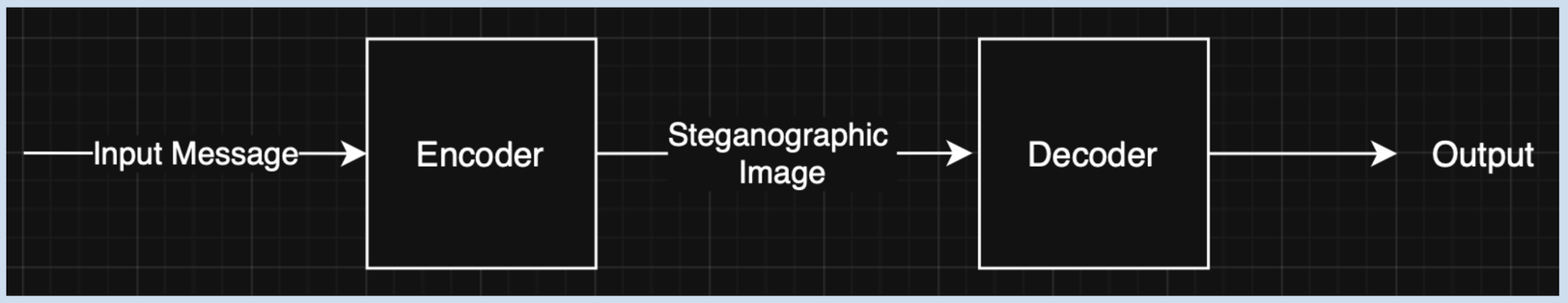
https://jwcneurasipjournals.springeropen.com/arti
022-02190-8

https://ieeexplore.ieee.org/document/9335027

https://emailsecurity.fortra.com/blog/how-to-defend-against-stegomalware

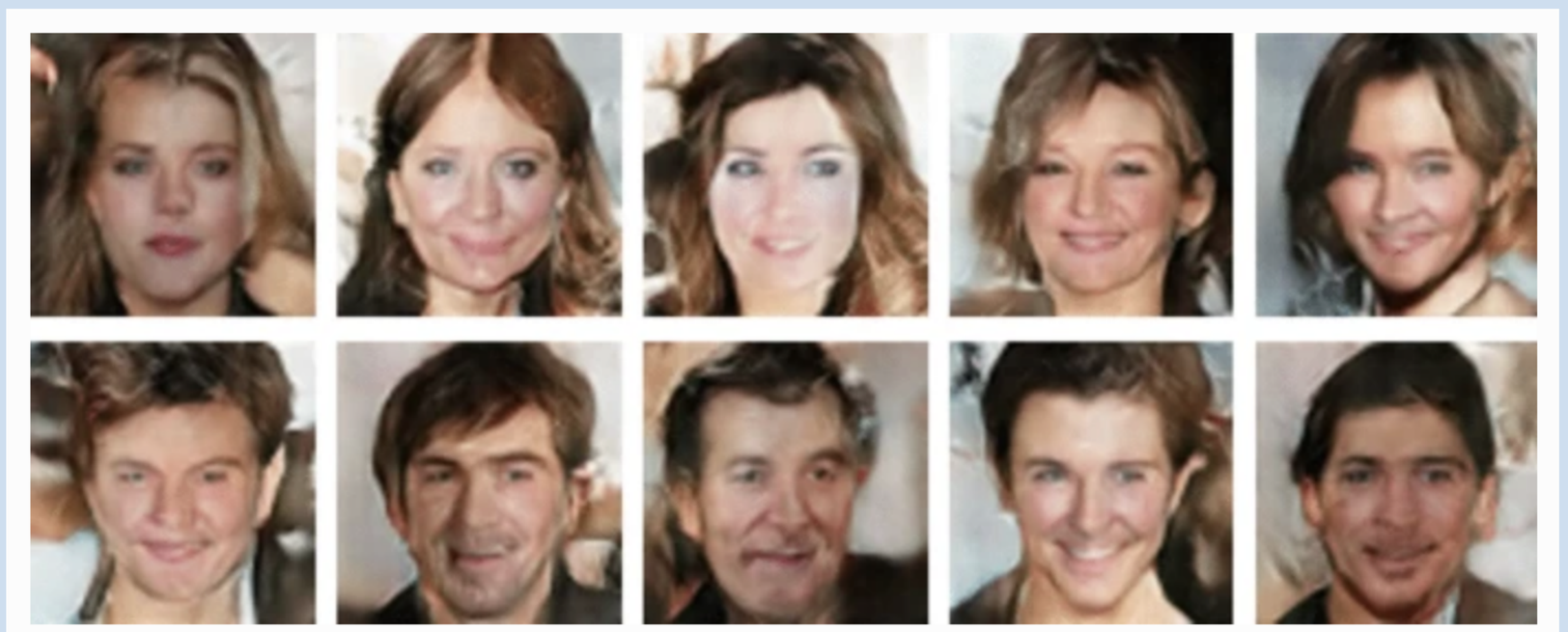
https://www.computer.org/csdl/magazine/sp/2022

General Structure



Generated Steganographic Images By Encoder

While training the Encoder model, we used the CelebA dataset, which contains faces of celebrities, so steganographic images are generated in the shape of human faces.



GAN Based Steganography Resistance Against Steganalyzer

Our method is resistant to stganalzers as seen in the diagram. There are currently no counter-techniques to generative steganography techniques.