

## Translation-based image steganography system utilizing autoencoder and CycleGAN

Thakwan Akram Jawad<sup>1</sup>, Jamshid Bagherzadeh Mohasefi<sup>2</sup>, Mohammed Salah Reda Abdelghany<sup>3</sup>

<sup>1</sup>Department of System and Control Engineering, College of Electronic Engineering, Ninevah University, Mosul, Iraq

<sup>2</sup>Department of Computer Engineering, Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran

<sup>3</sup>Department of Computers and Artificial Intelligence, Damietta University, Damietta, Egypt

### Article Info

#### Article history:

Received Aug 24, 2024

Revised Jun 16, 2025

Accepted Jul 10, 2025

#### Keywords:

Autoencoder

Coverless steganography

CycleGAN

Generative models

Steganography

### ABSTRACT

Traditional image steganography involves embedding secret information into a cover image, a process that requires modification of the carrier and potentially leaves detectable marks. This paper proposes a novel method of coverless image steganography based on generative models. Initially, a CycleGAN model is constructed and trained to learn the features of different image domains. Subsequently, an Autoencoder model is trained using two sets of images to achieve a precise one-to-one mapping. Once the models are trained, the autoencoder is used on both the sender and receiver sides to convert the cover image (also known as the stego image) into the secret image and vice versa. The CycleGAN model is then utilized to enhance the visual quality of the images generated by the autoencoder. Experimental results demonstrate that this method not only effectively secures secret information transmission but also improves efficiency and increases the capacity for information hiding compared to similar methods.

This is an open access article under the [CC BY-SA](#) license.



### Corresponding Author:

Jamshid Bagherzadeh Mohasefi

Department of Computer Engineering, Faculty of Electrical and Computer Engineering, Urmia University

11 km Serow Road, Urmia, West Azerbaijan, Iran

Email: j.bagherzadeh@urmia.ac.ir

## 1. INTRODUCTION

Steganography is the science and art of making the information invisible by using different concealing methods. Similar to cryptography, steganography tries to ensure the secrecy and safety of information, but unlike cryptography, steganography tries to hide the existence of this secret information [1]. The most well-known embedding algorithm is the least significant bit (LSB) algorithm. The idea of LSB is to encode the secret message into the LSB of a color channel in the cover image. Since the LSB algorithm manipulates the pixel independently of each other, it is vulnerable to steganalysis techniques and prone to detection [2].

Two crucial requirements of every traditional steganography system are the data and a carrier. A carrier is referred to the paper, image, video, or any multimedia that carries the secret data. The data is embedded into the carrier, and then transmitted to recipient. However, this method raises some issues, because the secret data is embedded into the cover, for instance by manipulating the pixels of the carrier image, attacker can detect the existence of a secret message in the image or any medium it is being transmitted with [1], [3]–[5].

Steganalysis, an emerging field of study parallel to steganography, is referred to the science of detecting the existence of hidden data in the cover file. The approaches utilized for steganalysis sometimes depend on the steganography algorithm(s) used to conceal the data [6]. In order to address the concerns with the conventional steganography, experts proposed coverless information hiding in 2014. Coverless

information hiding is referred to the natural carrier, which is compelled by the secret data. By sharing the mapping between certain features of the carrier and the secret data, sender and receiver can communicate secret information without changing or manipulating the medium also known as the cover [5].

Generative adversarial networks (GAN) are a variation of deep convolutional neural network (CNNs) put forth by Almahairi *et al.* [7]. A GAN is made of two deep networks generator and discriminator compete against each other in a zero-sum game to produce a valid output in GAN architecture. The generator network is trained in a way to produce similar images to its input and fool the discriminator. The discriminator is trained to find the fake images effectively [8]. Generally, a GAN model is made of two chief components. Which are discriminator and generator. A new network named the steganalyzer, with the purpose of checking if the input has any concealed data in it or not, is used in some methods within the image steganography context [8].

CycleGAN, abbreviated from cycle-consistent generative adversarial networks, is a generative model employed in computer vision and image synthesis. It was devised to facilitate unsupervised learning for image-to-image translation, eliminating the need for paired training data [9]. In the CycleGAN framework, a generator is trained to produce images in one domain based on images from another domain. Since there is no reliance on paired information, numerous potential mappings could be inferred. To constrain the multitude of possible mappings, CycleGAN is commonly trained with a cycle-consistency constraint. This constraint ensures a robust connection across domains by mandating that the transformation of an image from the source domain to the target domain and back to the source should yield the original image [7].

In our research, we will draw inspiration from novel concepts presented in the studies conducted prior to this research, specifically adopting feature mapping technique. However, a significant departure in our proposed steganography system is its fully-coverless nature. When compared to traditional steganographic methods, our coverless steganography solution exhibits several advantages:

- Imperceptibility: traditional methods often leave detectable traces, making them susceptible to steganalysis. Our method, however, produces images that are virtually indistinguishable from genuine images, significantly enhancing imperceptibility.
- Higher data capacity: by not embedding data directly into the image pixels, our method circumvents the capacity limitations of traditional techniques, allowing for more substantial information to be concealed.
- Efficiency: the use of autoencoders and CycleGAN allows for efficient data concealment and retrieval processes. This efficiency is critical for real-time applications and scenarios where computational resources may be limited.

This paper is structured as follows: the second section provides a comprehensive review of prior research in the field, establishing the context for this study. The third section details our proposed method, outlining its design and implementation. In the fourth section, we evaluate the effectiveness of our approach through two carefully designed experimental scenarios. Finally, the fifth section concludes the paper, summarizing key findings and implications of this research.

## 2. RELATED WORK

GAN are variations of deep CNNs [10], [11] put forth by Goodfellow *et al.* [12]. A GAN is made of two deep networks generator and discriminator compete against each other in a zero-sum game to produce a valid image in GAN architecture. The generator network is trained in a way to produce similar images to its input and fool the discriminator, and the discriminator is trained to find the fake images effectively [13]. Generally, a GAN model is made of two chief components. Which are discriminator and generator. A new network named the steganalyzer, with the purpose of checking if the input has any concealed data or not, is used in some methods within the image steganography context [13]. There are different variations of GAN being used in steganography tasks. To name some, CycleGAN, conditional generative adversarial networks (CGAN) [14], deep convolutional generative adversarial networks (DCGAN), and Wasserstein generative adversarial networks (WGAN) [15].

Hu *et al.* [16] proposed the DCGAN architecture used for steganography without embedding (SWE). Their proposed method eliminates the embedding process by generating the carrier image based on the noise vector which secret information was mapped to. In this method another network called the extractor network is required to re-acquire the secret information from the carrier image.

While some steganographic approaches utilize a single generative model, Li *et al.* [17] proposed a two-stage method with separate models (F and G) for cover image generation and secret image reconstruction. However, their work acknowledges challenges in reconstructing the secret image due to a lack of content information preserved during the initial cover image generation step by model F. To address this issue, the authors introduce a novel "content-consistency" extraction module within the cover image generation process.

Li *et al.* [18] incorporated a DCGAN architecture for the discriminator. In their study, they have built a method that transfers secret images between two domains. Their work proposes a two-stage generative model approach for secret image encryption. In the first stage, they embed the secret image within a public image from a different domain, resulting in a synthetic image. This synthetic image is then used as input for a generative model (F) to generate an encrypted image in another domain.

Zhang *et al.* [19] proposed a generative reversible data hiding (GRDH) approach utilizing image translation. Their method involves two stages. First, an image generator creates a realistic image that serves as input to a CycleGAN model. CycleGAN performs image-to-image translation, resulting in a stego-image that conceals the secret message. Notably, both the secret message and the original image can be recovered independently. A trained message extractor retrieves the hidden message, while the inverse of the image translation process recovers the original image.

Duan and Song *et al.* [13] introduced a novel coverless image information hiding method utilizing generative model database. They propose transmitting a newly generated image, independent from the secret image, that can be decoded back to the original secret image by the receiver using their generative model database. This method offers improved security by transmitting an uninformative image. They utilize WGAN model to achieve coverless image information hiding. Instead of directly transmitting the secret image, they train the WGAN on the secret image (replacing the usual random noise input). This WGAN can then generate a new, independent "meaning-normal" image that doesn't contain the secret information. This produced picture is transmitted to the receiver, which uses WGAN model to regenerate the original secret image. However, this method confronts scalability challenges. The current approach imposes storing and managing two separate WGAN models for each sender-receiver set. This can become bulky for large-scale deployments or scenarios with frequent communication partner changes. Additionally, the computational cost linked with training and maintaining these models could be a limiting issue.

StarGAN has been introduced by Choi *et al.* [20], which is a GAN that has the ability to multi-domain image transformation that can be used to achieve facial attribute manipulation within the cover image. In this network, the training process comprises both the generator network (G) and a discriminator network (D); the generator takes a target domain tag that represents the desired attributes and the input image. After that, a new image will be generated to incorporate these specified attributes. For the discriminator network, it has two functions where they differentiate between real and fake images in addition to recognizing the domain (attributes) associated with the input image. A combination of loss functions will be held by StarGAN to achieve these goals, this includes adversarial, classification, and reconstruction losses.

By employing an original form of GAN architecture, Shi *et al.* [21] deviates from existing techniques in their proposed method that consists of one generative network and two discriminative networks. The role of the generative network is to prioritize the visual quality of the stego images, while the role of the discriminative networks is to assess their suitability for information hiding. The authors declare considerable improvements in the speed of convergence, stability of the training, and the quality of the image. Also, they used an advanced steganalysis network within the discriminative structure; this enables better evaluation of the generated images' performance. In this work, three aspects have been prioritized which are perceptibility, security, and diversity. To achieve a high-quality stego image, they used a WGAN instead of the commonly used DCGAN, which leads to faster training and superior visual quality.

### 3. METHOD

The proposed method involves designing two advanced artificial neural networks (ANNs) as core components of the system. The objective is to create a one-to-one mapping between two distinct image sets. Ideally, the model should perfectly transform an input image (A) into a completely different output image (B). This image B can belong to the same domain as image A or a different domain altogether. Crucially, image B should exhibit no discernible traces of the information originally contained within image A. To achieve this, the system incorporates two generative neural networks:

- Autoencoder: trained using two sets of images to map them one-to-one. It is used on both the sender and receiver sides to convert the cover image (stego image) to the secret image and vice versa.
- CycleGAN: enhances the visual quality of the images generated by the Autoencoder. The CycleGAN model is trained to learn the features of the images' domains, ensuring high-quality image translation between domains.

#### 3.1. Datasets

The datasets employed for this study include images from the CelebA and WikiArt datasets. These datasets provide a diverse range of images necessary for training the generative models effectively. The CelebA dataset includes images of celebrities, while the WikiArt dataset encompasses various styles of

art [22], [23]. To evaluate the model's performance in a comprehensive way, both datasets are split into training and testing sets.

### 3.2. Architectures

#### 3.2.1. Autoencoder

The idea of designing the encoder is to reduce the spatial dimensions of the input image in a gradual way while expanding the depth of the feature maps and capturing basic image features. In the beginning of the structure, there are three layers, a Conv2D layer that uses 32 filters with rectified linear unit (ReLU) activation, a MaxPooling2D layer that uses pool size of  $2 \times 2$ , and a dense layer comprising 64 units with ReLU activation. After that, the process will be repeated with Conv2D layer of 64 filters, another layer of MaxPooling2D, and a dense layer that has 128 units. At the final stage, the encoder has a Conv2D layer that have 128 filters, a MaxPooling2D layer, and a dense layer that contains 256 units. All use ReLU activation to guarantee non-linearity and robust feature extraction.

The function of the decoder is to refine the image from the encoded latent illustration, this done by reversing the dimensionality reduction. It starts with the dense layer that has  $8 \times 8 \times 128$  units, a Conv2DTranspose layer with 256 filters of size  $3 \times 3$  and a stride of 2 and ReLU activation, and a dense layer with 64 units. Then, it continues with a Conv2DTranspose layer that has 28 filters and a stride of 2, followed by a dense layer that has 32 units. The final step have Conv2DTranspose layer that has 64 filters, a stride of 2, and a concluding Conv2D output layer that has with 3 filters of size  $3 \times 3$ . It employs sigmoid activation to produce the final output image.

#### 3.2.2. CycleGAN

The architecture of the CycleGAN generator includes three main modules which are downsampling, residual blocks, and upsampling. The phase of downsampling starts with a Conv2D layer that has 64 filters, a  $4 \times 4$  kernel, and a stride of 2. To reduce the spatial dimensions by half, ReLU activation will be added. To further halve the dimensions, the Conv2D layer will be used. This layer has 128 filters while following the same kernel and stride to increase feature depth. To capture more complex features while continuing the spatial dimension reduction, a Conv2D layer with 256 filters,  $4 \times 4$  kernel, and stride of 2 has been used. To maintain the network's identity mappings and training constancy, outstanding blocks will be used as their use is crucial. The six outstanding blocks comprising a Conv2D layer with 256 filters, a  $3 \times 3$  kernel, and same padding that followed by ReLU activation. Another Conv2D layer that has 256 filters and same padding follows that the input and output summed to form the residual connection, promoting stable training [24].

The upsampling stage starts with a Conv2DTranspose layer that has 256 filters, a  $4 \times 4$  kernel, and a stride of 2. It will be ctivated by ReLU, doubling the spatial dimensions, followed by a Conv2DTranspose layer that has 128 filters and the same kernel and stride. To further upsamples the dimensions, another Conv2DTranspose layer with 64 filters and similar settings. Finally, the output layer is a Conv2DTranspose layer with 3 filters, a  $4 \times 4$  kernel, and a stride of 2. It will be followed by Sigmoid activation to constrain output pixel values between 0 and 1 [25], [26].

For the discriminator, it has been designed as a PatchGAN that focuses on local image patches rather than the entire image, maintaining high-frequency details and textures. It has multiple Conv2D layers with filters and strides of 2, each followed by LeakyReLU activations. It progressively reducing the spatial dimensions while capturing more complex features. The final layer outputs a single-channel feature map that represents the authenticity of each image patch, promoting detailed and realistic image generation [9], [27].

## 4. RESULTS AND DISCUSSION

The functionality of our proposed method is evaluated through various performance metrics, including structural similarity index (SSIM) [28], mean squared error (MSE), and peak signal-to-noise ratio (PSNR). To comprehensively evaluate our proposed system, we have defined two distinct scenarios. These scenarios are designed to test the system's performance in different contexts and with different types of images.

### 4.1. Training CycleGAN

We selected 50 images each from the CelebA and action painting style in the WikiArt dataset to train the CycleGAN model, aiming to represent two distinct visual domains and thus provide a robust test for our model's generalization capabilities. The results and visualizations include recorded training losses for both the generators and discriminators over the 500 epochs. The plot in Figure 1 illustrates the training dynamics, specifically the loss curves for both generators (Gen AtoB and Gen BtoA) as they learn to translate images from one domain to another. The CycleGAN model was trained for 500 epochs to learn the mappings between these two domains, capturing and translating the unique features and styles of each domain, as

shown in Figures 2, CelebA to WikiArt as shown in Figure 2(a), and WikiArt to CelebA as shown in Figure 2(b). This training enhances the visual imperceptibility of the generated images.

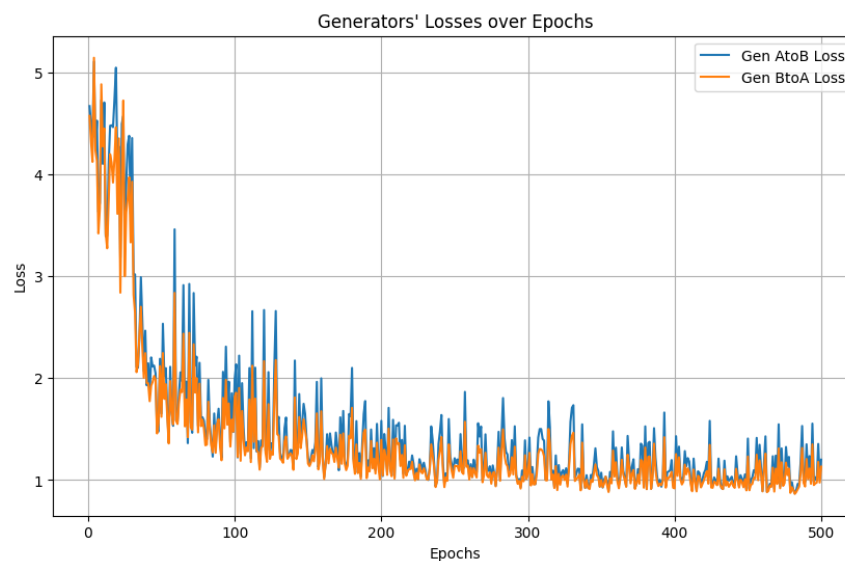


Figure 1. Generators' training loss over epochs

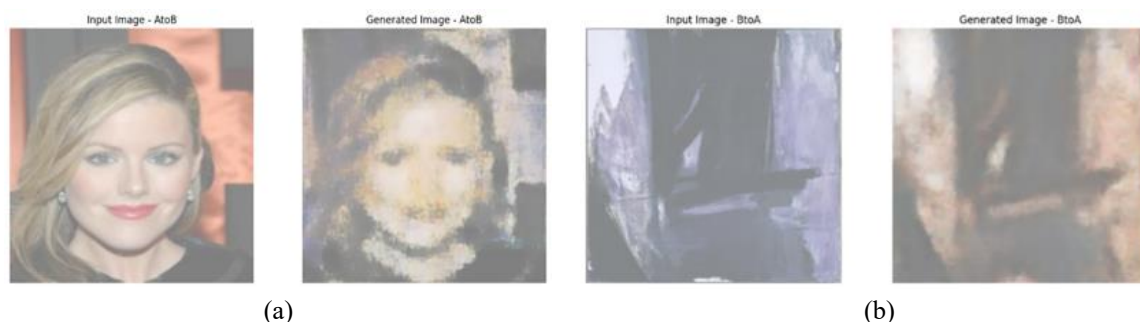


Figure 2. Examples of CycleGAN domain-to-domain translation after the model is trained of (a) CelebA to WikiArt and (b) WikiArt to CelebA

#### 4.2. Scenario 1: same domain

In this scenario, both the secret image and the stego image belong to the same domain, sharing similar characteristics such as style, content, and visual features. Evaluating the system under this condition helps us understand its performance when the images have a high degree of similarity. We selected 20 images from the CelebA dataset, featuring both male and female subjects with diverse backgrounds, to test the system's ability to map and translate facial images. These images, with relatively simpler and more uniform visual features compared to artistic images, provide an appropriate test case to assess the system's efficacy in handling facial imagery. In the initial phase of our evaluation, we focused on the first scenario, where both the stego-image and the secret image originate from the same domain. This scenario is particularly significant as it allows us to assess the performance of our model under conditions where the features of the images are highly similar.

The similarity in features between the stego-image and the secret image simplifies the learning process for our model, thereby serving as a baseline for evaluating its effectiveness. In this evaluation, we used 20 images from the CelebA dataset. These images featured both male and female subjects with varied backgrounds. We randomly selected 20 images and subsequently divided them into two distinct sets, labeled A and B, as illustrated in the Figure 3. This process ensures a balanced and representative sample, which is

essential for evaluation of our proposed method. Figure 4 illustrates the autoencoder's training loss trajectory over 2000 epochs, reflecting the optimization process. Following this, Figure 5 evaluates the model's performance by showing original images in Figure 5(a), their generated counterparts in Figure 5(b), and the target secret images in Figure 5(c), highlighting the model's reconstruction accuracy.



Figure 3. The images of two sets that will be mapped correspondingly

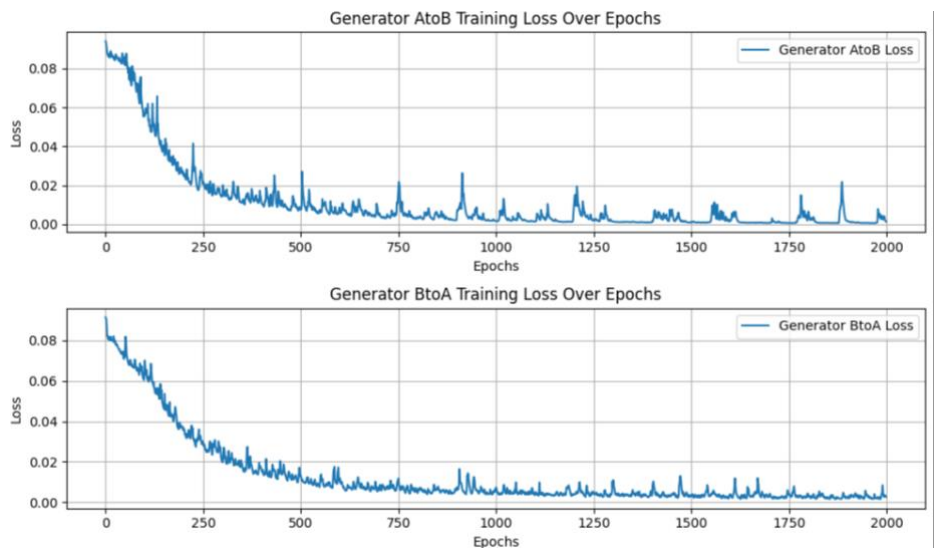


Figure 4. The training loss over epochs during the training of the autoencoder on the CelebA dataset

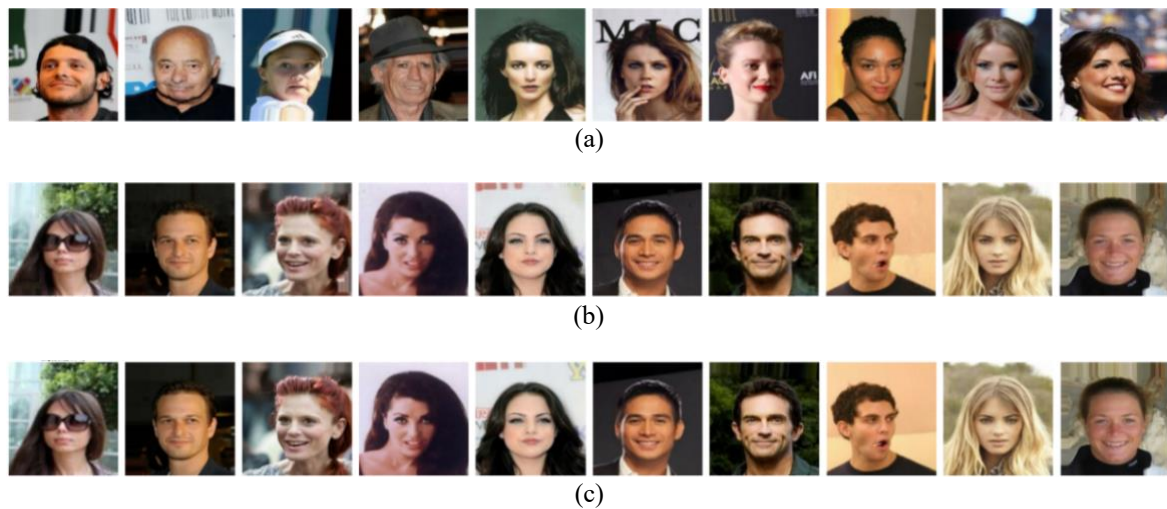


Figure 5. Autoencoder model evaluation of (a) input: original image, (b) output: generated image, and (c) target: secret image



As shown in Table 1, the analysis of the CelebA-to-CelebA translation performance for the autoencoder model reveals noteworthy findings from the comparative assessment of SSIM, MSE, and PSNR metrics across different image sets. In the AtoB model, the results demonstrate consistently high SSIM values, peaking at 0.908185 for image set 4, indicating strong preservation of structural integrity during translation. Additionally, this model exhibits low MSE values, particularly notable in image set 4 with an MSE of 0.000360, and a corresponding PSNR of 34.434243, signifying minimal reconstruction error and high-quality output images. These results collectively highlight the robustness of the autoencoder model in maintaining image quality, while also suggesting areas for improvement in achieving consistent performance across different image sets.

Table 1. Evaluation metrics for generated images in Figure 5

Image set	SSIM	MSE	PSNR
Image set 1	0.859619	0.003551	24.496793
Image set 2	0.825956	0.000952	30.212155
Image set 3	0.866684	0.002285	26.411011
Image set 4	0.908185	0.000360	34.434243
Image set 5	0.864702	0.000943	30.256875
Image set 6	0.855163	0.000631	32.002772
Image set 7	0.759118	0.001630	27.877301
Image set 8	0.904223	0.000509	32.932824
Image set 9	0.895943	0.000760	31.191168
Image set 10	0.863341	0.000718	31.439247

4.3. Scenario 2: different domains

The secret image and the stego image come from different domains in this scenario to present a more challenging setup due to their dissimilar characteristics in style, content, and visual features. The evaluation of this system under these conditions helps to understand its robustness and versatility. Secret images from the CelebA dataset and stego images from the WikiArt dataset have been used to test the system's ability to map and translate facial images to artistic images. The objective was ensuring the facial features were preserving while maintaining the artistic integrity of the stego images. In the second phase of the evaluation, it has been focused on the second scenario, where the stego-image and the secret image have been originated from different domains. This scenario is crucial as it gives the ability to evaluate the performance of the model under conditions where the features of the images are highly different and unsimilar. It is important to mention that the difference in features between the stego-image and the secret image complicates the learning process for our model.

As shown in Figure 6, twenty images were randomly selected for evaluation. These images have been selected from the WikiArt and CelebA datasets equally. The reason for choosing these images is to cover a diverse range of artistic styles and intricate visual and facial features to assess the ability of the system to translate and map images across different domains. In Figure 7, the autoencoder's training loss over epochs for both datasets, this highlights the optimization process and the model's convergence during training. While Figure 8 is used to evaluate the autoencoder's performance by showcasing the original images in Figure 8(a), then the generated images in Figure 8(b), and the target secret images in Figure 8(c), demonstrating the model's capacity for effective reconstruction and domain mapping.

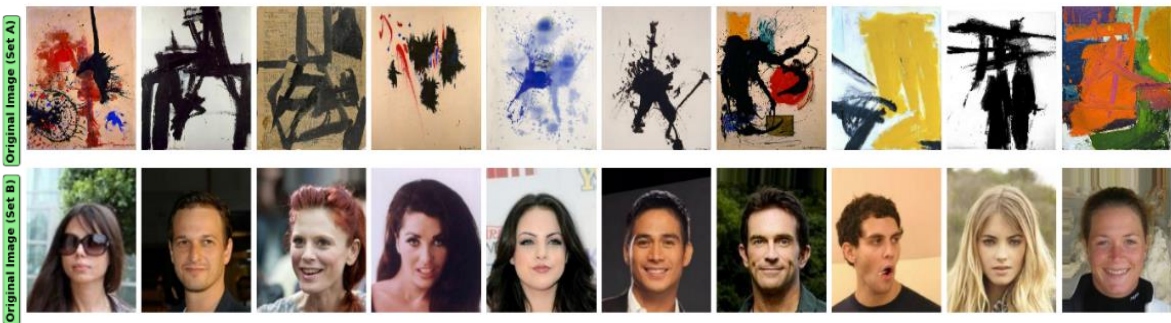


Figure 6. The images of two sets that will be mapped correspondingly

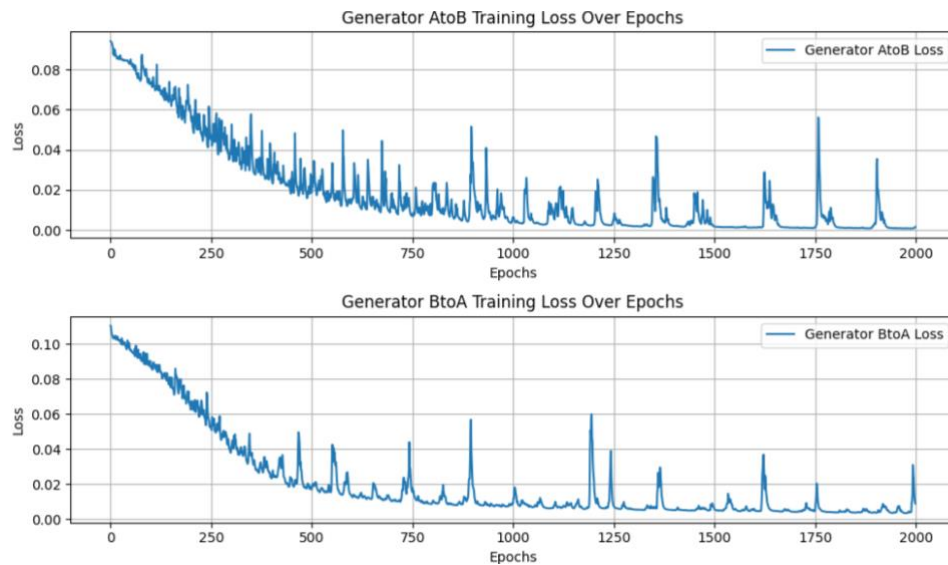


Figure 7. The training loss over epochs during the training of the autoencoder on the WikiArt/CelebA datasets

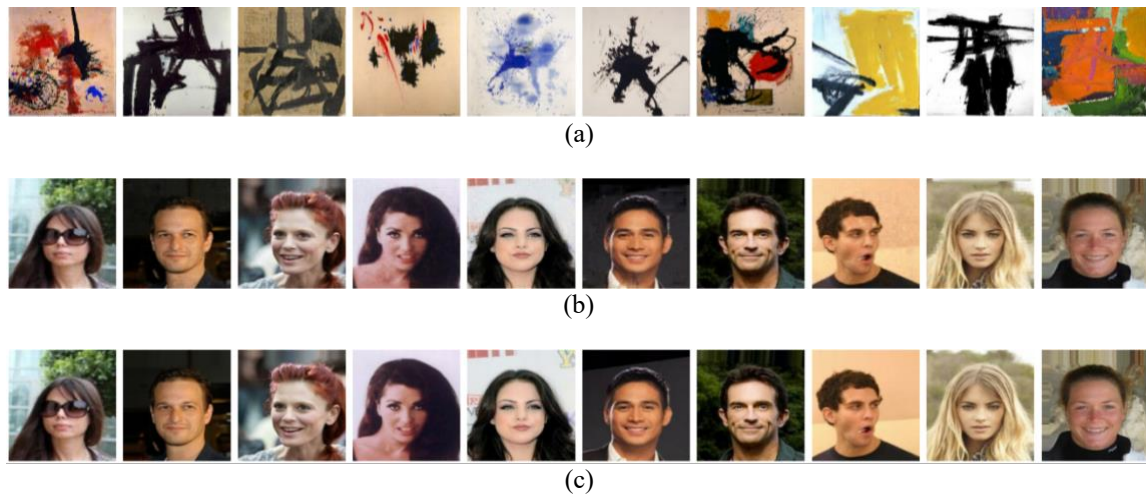


Figure 8. Autoencoder model evaluation of (a) input: original image, (b) output: generated image, and (c) target: secret image

In addition, Table 2 contributes a quantitative assessment of the autoencoder's performance. It presents evaluation metrics for the generated images displayed in Figure 8. These metrics confirm the effectiveness of the model to maintain high accuracy and quality in the reconstructed images.

Table 2. Evaluation metrics for generated images in Figure 8

Image set	SSIM	MSE	PSNR
Image set 1	0.596109	0.009241	20.342598
Image set 2	0.649340	0.007551	21.219742
Image set 3	0.493413	0.005901	22.290604
Image set 4	0.697044	0.005046	22.970548
Image set 5	0.628170	0.004263	23.703005
Image set 6	0.719750	0.004531	23.437798
Image set 7	0.620793	0.006865	21.633631
Image set 8	0.673014	0.007311	21.360182
Image set 9	0.707038	0.007211	21.419772
Image set 10	0.476001	0.007411	21.301239



#### 4.3.1. Applying CycleGAN

As the previous results demonstrated, the autoencoder's performance on image pairs 1, 3, and 10 wasn't good enough. Therefore, in this section, we are going to apply our trained CycleGAN to the images that were outputted from the autoencoder to enhance them. We will then compare these enhanced images both metrically and visually to their respective target images. Figure 9 evaluates the CycleGAN model, showcasing its capability to enhance the visual quality of the autoencoder's output. Figure 9(a) displays the original image, Figure 9(b) presents the improved output generated by the CycleGAN, and Figure 9(c) illustrates the target secret image, emphasizing the model's contribution to refining output quality. Table 3 provides a comparative analysis of evaluation metrics for both the generated images from the autoencoder and the enhanced images produced by the CycleGAN. These metrics illustrate the improvements in visual quality and fidelity.

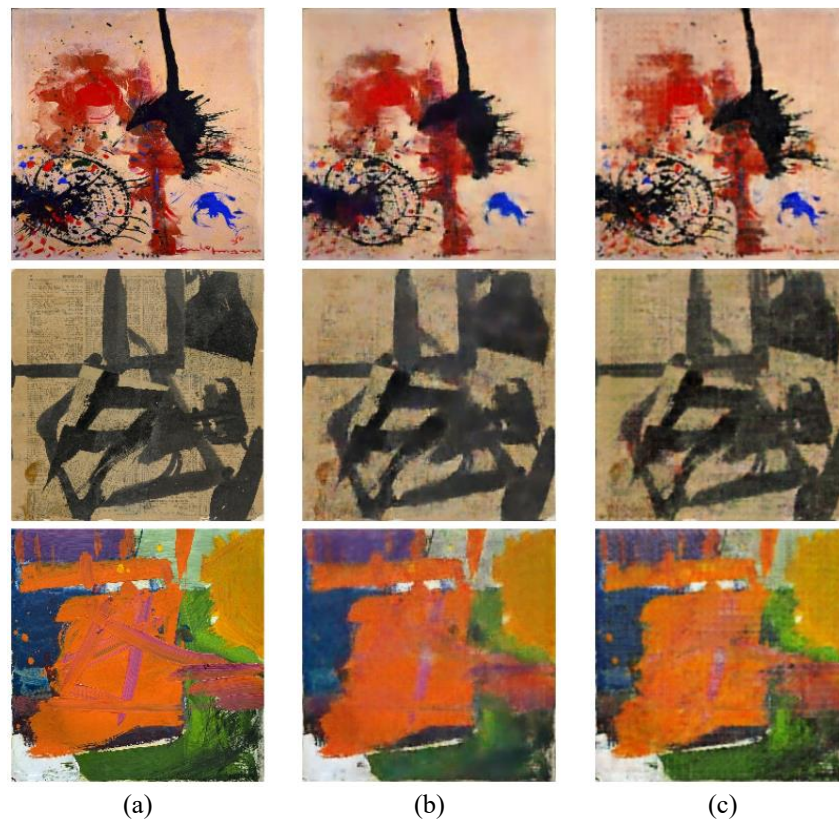


Figure 9. CycleGAN model evaluation of (a) input: original image, (b) output: generated image, and (c) target: secret image

Table 3. Evaluation metrics for the generated images and enhanced images

Image set	Model	SSIM	MSE	PSNR
1	Autoencoder	0.596109	0.009241	20.342598
	CycleGAN	0.604278	0.009573	20.189400
3	Autoencoder	0.493413	0.005901	22.290604
	CycleGAN	0.505235	0.005854	22.325549
10	Autoencoder	0.476001	0.007411	21.301239
	CycleGAN	0.481077	0.007629	21.175290

#### 4.4. Comparison of steganography techniques

To evaluate the effectiveness of various steganographic techniques, we conducted a comparative analysis based on key factors: cover image size, secret image size, security, visual fidelity, and versatility. The cover/stego image size refers to the resolution of the image used for embedding the secret data in traditional steganographic methods and also represents the resolution of the transmitted image required to extract the hidden data. Secret image size denotes the resolution of the data being concealed within the cover

image. Security measures the resistance of the generated images against detection by steganalysis tools, ensuring that the hidden data remains imperceptible. Versatility assesses the model's adaptability to different image domains, while visual fidelity reflects the quality and realism of the generated images. In Table 4, we provide a detailed comparison of our method with existing techniques.

Table 4. Comparison between our proposed method and previous methods

Method	Technique	Cover/stego image size	Secret image size	Security	Visual fidelity	Versatility
Proposed method	Autoencoder+CycleGAN	256×256×3	256×256×3	High	Moderate	Very high
Rehman's method [29]	Encoder-decoder	32×32×3	32×32×1	Moderate	High	High
Zhang's method [30]	GAN	256×256×3	256×256×1	Moderate	High	High

#### 4.5. Limitations

In this study, one significant limitation is the scarcity of established baseline models for comparative evaluation. Since this field is relatively nascent, there is a noticeable lack of standardized datasets and performance metrics that can serve as benchmarks for assessing the effectiveness of our model. This gap presents challenges in quantitatively evaluating our approach against a diverse array of methods. Moving forward, the establishment of standardized benchmarks and evaluation frameworks will be critical in facilitating more robust comparisons and fostering advancements across various models in this domain. Additionally, while the CycleGAN effectively enhances the visual quality of the generated images, it does not achieve complete enhancement, leaving some room for improvement in terms of image fidelity. Future work could explore optimizing the training process for the autoencoder and improving the CycleGAN's architecture to achieve higher levels of enhancement, thereby addressing these challenges.

## 5. CONCLUSION

This study demonstrates the feasibility of using generative models, particularly CycleGAN and autoencoders, for coverless image steganography, achieving significant advancements in security, data capacity, and visual quality. Unlike traditional methods that modify carrier images, our approach generates undetectable stego images. The integration of CycleGAN ensures high visual fidelity, while autoencoders enable effective data concealment and recovery. These findings highlight the robustness, efficiency, and applicability of the proposed method across various domains, paving the way for more secure and practical solutions in steganography. Future research will focus on overcoming limitations and further optimizing the method for real-world applications.

## FUNDING INFORMATION

Authors state no funding involved.

## AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Thakwan Akram Jawad	✓	✓		✓		✓	✓	✓	✓					✓
Jamshid Bagherzadeh	✓				✓					✓	✓	✓	✓	
Mohasefi														
Mohammed Salah Reda Abdelghany			✓	✓		✓			✓	✓				

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

## CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.





## DATA AVAILABILITY

The data that supports the findings of this study are openly available in Kaggle repository as non-commercial research purposes only at <https://www.kaggle.com/datasets/jessicali9530/celeba-dataset> and <https://www.kaggle.com/datasets/steubk/wikiart>.





## REFERENCES

- [1] N. F. Johnson and S. Jajodia, "Exploring steganography: seeing the unseen," *Computer*, vol. 31, no. 2, pp. 26–34, Feb. 1998, doi: 10.1109/MC.1998.4655281.
- [2] D. Volkhonskiy, I. Nazarov, and E. Burnaev, "Steganographic generative adversarial networks," *Twelfth International Conference on Machine Vision (ICMV 2019)*, 2019, doi: 10.1117/12.2559429.
- [3] A. Cheddad, J. Condell, K. Curran, and P. M. Kevitt, "Digital image steganography: survey and analysis of current methods," *Signal Processing*, vol. 90, no. 3, pp. 727–752, 2010, doi: 10.1016/j.sigpro.2009.08.010.
- [4] R. J. Anderson and F. A. P. Petitcolas, "On the limits of steganography," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 4, pp. 474–481, 1998, doi: 10.1109/49.668971.
- [5] R. Meng, Z. Zhou, Q. Cui, X. Sun, and C. Yuan, "A novel steganography scheme combining coverless information hiding and steganography," *Journal of Information Hiding and Privacy Protection*, vol. 1, no. 1, pp. 43–48, 2019, doi: 10.32604/jihpp.2019.05797.
- [6] N. Meghanathan and L. Nayak, "Steganalysis algorithms for detecting the hidden information in image, audio and video cover media," *International Journal of Network Security & Its Application*, vol. 2, no. 1, 2010.
- [7] A. Almahairi, S. Rajeswar, A. Sordoni, P. Bachman, and A. Courville, "Augmented cyclegan: learning many-to-many mappings from unpaired data," *arXiv-Computer Science*, pp. 1–10, 2018.
- [8] N. Subramanian, O. Elharrouss, S. Al-Maadeed, and A. Bouridane, "Image steganography: a review of the recent advances," *IEEE Access*, vol. 9, pp. 23409–23423, 2021, doi: 10.1109/ACCESS.2021.3053998.
- [9] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2223–2232, 2017, doi: 10.1109/ICCV.2017.244.
- [10] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.
- [11] K. Fukushima, "Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, no. 4, pp. 193–202, Apr. 1980, doi: 10.1007/BF00344251.
- [12] I. Goodfellow *et al.*, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: 10.1145/3422622.
- [13] X. Duan and H. Song, "Coverless information hiding based on generative model," *arXiv-Computer Science*, pp. 1–4, 2018.
- [14] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv-Computer Science*, pp. 1–7, 2014.
- [15] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," *arXiv-Statistics*, pp. 1–32, 2017.
- [16] D. Hu, L. Wang, W. Jiang, S. Zheng, and B. Li, "A novel image steganography method via deep convolutional generative adversarial networks," *IEEE Access*, vol. 6, pp. 38303–38314, 2018, doi: 10.1109/ACCESS.2018.2852771.
- [17] Q. Li, X. Wang, X. Wang, and Y. Shi, "CCCIH: content-consistency coverless information hiding method based on generative models," *Neural Processing Letters*, vol. 53, no. 6, pp. 4037–4046, 2021, doi: 10.1007/s11063-021-10582-y.
- [18] Q. Li, X. Wang, X. Wang, B. Ma, C. Wang, and Y. Shi, "An encrypted coverless information hiding method based on generative models," *Information Sciences*, vol. 553, pp. 19–30, 2021, doi: 10.1016/j.ins.2020.12.002.
- [19] Z. Zhang, G. Fu, F. Di, C. Li, and J. Liu, "Generative reversible data hiding by image-to-image translation via GANs," *Security and Communication Networks*, no. 1, 2019, doi: 10.1155/2019/4932782.
- [20] Y. Choi, M. Choi, M. Kim, J. W. Ha, S. Kim, and J. Choo, "StarGAN: unified generative adversarial networks for multi-domain image-to-image translation," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 8789–8797, 2018, doi: 10.1109/CVPR.2018.00916.
- [21] H. Shi, J. Dong, W. Wang, Y. Qian, and X. Zhang, "SSGAN: secure steganography based on generative adversarial networks," in *Advances in Multimedia Information Processing*, Cham, Switzerland: Springer, 2018, pp. 534–544.
- [22] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," *2015 IEEE International Conference on Computer Vision*, pp. 3730–3738, 2015, doi: 10.1109/ICCV.2015.425.
- [23] W. R. Tan, C. S. Chan, H. E. Aguirre, and K. Tanaka, "Improved ArtGAN for conditional synthesis of natural image and artwork," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 394–409, 2019, doi: 10.1109/TIP.2018.2866698.
- [24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
- [25] M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus, "Deconvolutional networks," *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2528–2535, 2010, doi: 10.1109/CVPR.2010.5539957.
- [26] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," *arXiv-Statistics*, pp. 1–31, 2018.
- [27] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," *arXiv-Computer Science*, pp. 1–5, 2015.
- [28] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004, doi: 10.1109/TIP.2003.819861.
- [29] A. Rehman, R. Rahim, S. Nadeem, and S. ul Hussain, "End-to-end trained CNN encoder-decoder networks for image steganography," in *Computer Vision – ECCV 2018 Workshops*. Cham, Switzerland: Springer, 2019, pp. 723–729, doi: 10.1007/978-3-030-11018-5\_64.
- [30] R. Zhang, S. Dong, and J. Liu, "Invisible steganography via generative adversarial networks," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 8559–8575, 2019, doi: 10.1007/s11042-018-6951-z.





**BIOGRAPHIES OF AUTHORS**

**Thakwan Akram Jawad**     received his Master of Science in Computer Engineering from Northern Cyprus (Eastern Mediterranean University) in Türkiye. He is assistant lecturer at Department of System and Control Engineering, College of Electronic Engineering, Ninevah University, Mosul, Iraq. He can be contacted at email: [thakwan.jawad@uoninevah.edu.iq](mailto:thakwan.jawad@uoninevah.edu.iq).



**Jamshid Bagherzadeh Mohasefi**     holds a Ph.D. in Computer Science from Indian Institute of Technology, India. He is currently Professor of the Department of Computer Science at the Urmia University. His research topics include artificial intelligence, machine learning, and network security. His research has been funded by the Urmia University, Iranian Telecom Ministry, and Iranian Ministry of Science, Research, and Technology. He can be contacted at email: [j.bagherzadeh@urmia.ac.ir](mailto:j.bagherzadeh@urmia.ac.ir).



**Mohammed Salah Reda Abdelghany**     holds a Ph.D. in Computer Science from the Faculty of Science at Damietta University, Egypt. He also earned a master's degree and a pre-master diploma in Computer Science from the Faculty of Science at Minia University. Currently, he is affiliated with the Faculty of Computers and Artificial Intelligence at Damietta University. His research interests span across various aspects of computer science and artificial intelligence. He can be contacted at email: [msr\\_cs\\_87@fcis.bsu.edu.eg](mailto:msr_cs_87@fcis.bsu.edu.eg).