FreqStegaGAN : Text Hiding in Image Frequency Bands via Critic-Regulated Adversarial Learning

Deep Piyushkumar Modi
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology (VIT), Vellore
Vellore, Tamil Nadu, 632014

Shantanu Anant Gaikwad
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology (VIT), Vellore
Vellore, Tamil Nadu, 632014

Skund Agarwal
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology (VIT), Vellore
Vellore, Tamil Nadu, 632014

Dr. Saravanan R
Prof. Higher Academic Grade
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology (VIT), Vellore
Vellore, Tamil Nadu, 632014

ABSTRACT

Steganography plays a crucial role in secure communication by embedding hidden messages in images. Traditional deep-learning-based approaches often operate in the spatial domain, leading to limitations in embedding capacity and visual quality. To address these issues, we propose FreqStegaGAN, a high-capacity image steganography model that leverages frequency-domain analysis. Our approach utilizes an Adaptive Frequency Channel Attention Network (AFcaNet), which dynamically assigns weights to frequency components and enhances feature extraction in a more meaningful way. The model utilizes a generative adversarial network (GAN) framework comprising an encoder, decoder, and a critic. The experimental results demonstrate that FreqStegaGAN outperforms prior methods in terms of embedding capacity, stego-image quality, and decoding accuracy.

I. INTRODUCTION

The demand for secure communication is growing, necessitating the development of advanced steganographic techniques. Image steganography aims to conceal messages within images in an imperceptible manner [6]. Early work in frequency-domain steganography employed genetic algorithms to optimize embedding locations, achieving good imperceptibility but limited capacity [1]. More recent deep-learning-based approaches have begun to exploit frequency-domain transforms alongside neural networks to improve both embedding capacity and robustness, for example by integrating attention mechanisms that prioritize important frequency bands [10], [3]. FreqStegaGAN combines these ideas by leveraging an adaptive frequency-domain channel attention mechanism to extract and weight frequency features efficiently [7],[8],[9], ensuring higher embedding capacity while preserving image quality.

The architecture of FreqStegaGAN is composed of three major components: encoder, decoder, and critic network. The encoder is responsible for embedding a secret message into the cover image, ensuring minimal visual distortion. It comprises convolutional layers, AFcaNet blocks for frequency attention, and dense connections for efficient feature propagation. The decoder reconstructs the hidden message from the generated stego-image by leveraging deep convolutional layers and AFcaNet to accurately recover the frequency-domain features. Finally, the critic network, inspired by Wasserstein GANs, guides the training process to improve the realism and imperceptibility of stego-images, thereby ensuring

that the generated images are visually indistinguishable from the original cover images [8],[9].

By integrating AFcaNet into the encoder and decoder, FreqStegaGAN enhances feature extraction in the frequency domain, allowing more effective message embedding without sacrificing visual quality. This approach overcomes limitations found in pure spatial-domain steganography and also improves upon earlier frequency-domain methods based solely on handcrafted optimization [1], [3],[6],[7].

II. METHODOLOGY

2.1 Adaptive Frequency Channel Attention Network (AFcaNet)

AFcaNet is an essential component of FreqStegaGAN, designed to optimize feature extraction in the frequency domain by assigning different levels of importance to different frequency components [7]. Unlike conventional convolutional networks that operate in the spatial domain, AFcaNet ensures that significant frequency features are enhanced while reducing distortions in less important regions [10],[6].

Purpose of AFcaNet

The Adaptive Frequency Channel Attention Network (AFcaNet) is designed to enhance image feature extraction by utilizing the frequency-domain characteristics of image data. Instead of relying purely on spatial features, it leverages Discrete Cosine Transform (DCT) to analyze and process frequency components of an image, ensuring better feature selection and representation [10],[7].

Why Use DCT in AFcaNet?

DCT is widely used in image compression and analysis because it efficiently represents signal energy in a few coefficients. The low-frequency components contain most of the image's structural information, while high-frequency components capture finer details. AFcaNet selects and enhances these frequency components dynamically to improve the learning process in neural networks [10],[7].

Working Mechanism :

AFcaNet follows the following sequential operations :

(a) Input Processing :

- The input x has dimensions (batch, C, H, W) where:
  - batch (B): Number of images in the batch.
  - C: Number of image channels (e.g., 3 for RGB).
  - H, W: Height and width of the image.

(b) Per-Channel Frequency Analysis using DCT:

- Each channel of the image is extracted and processed separately.
- The 2D Discrete Cosine Transform (DCT) was applied to obtain frequency coefficients:

$$DCT_{i,j}^{(c)} = \sum_{x=0}^{H-1}\sum_{y=0}^{W-1} X_{x,y}^{(c)} \cos\left(\frac{\pi i}{H}(x+0.5)\right)\cos\left(\frac{\pi j}{W}(y+0.5)\right)$$

- The top-left 7×7 block of DCT coefficients was selected because low-frequency components contain crucial information [10],[7].

(c) Feature Transformation :

- The 49 frequency components are flattened into vectors.
- A fully connected layer (fc1) processes 49 DCT coefficients, followed by ReLU activation to extract essential features.

(d) Weight Calculation :

- The outputs from all channels were concatenated to form a vector of size (B,C).
- Another fully connected layer (fc2) was applied, followed by a sigmoid function to normalize the weights to the range [0,1].
- The obtained weight vector W was reshaped to (B, C, 1, 1).

(e) Feature Recalibration :

- The original feature map x is multiplied channel-wise by the learned weight matrix to enhance or suppress the features dynamically.

2.2.Encoder

The Encoder takes a cover image C and a message M and embeds M into C to generate a stego image S. The primary goal is to make S visually similar to C, while ensuring that M is securely hidden [8],[9].

Why is the Encoder Used?

Traditional encoders only use spatial feature extraction, but here we integrate frequency-based attention

(AFcaNet) to select meaningful frequency features for steganographic embedding [10],[7],[9]. Earlier work showed that frequency-domain embedding guided by genetic algorithms could optimize imperceptibility, but capacity remained limited [1].

A dense feature concatenation strategy ensures that information from multiple layers is retained. Channel-wise attention mechanisms further enhance important features while suppressing irrelevant ones.

Working Mechanism of the Encoder :

The encoder comprises multiple convolutional blocks, frequency-based attention, and channel-wise attention.

Step 1 : Initial Feature Extraction

- The input cover image C (Shape: B × 3 × H × W) passes through the ConvBlock (3→32 channels) [8].
- It then undergoes AFcaNet processing to extract frequency-domain features [10]. Channel-wise attention was applied using adaptive average pooling to dynamically adjust feature importance [6].

Step 2 : Message Fusion and Feature Expansion

- The secret message M (Shape: B × D × H × W) was concatenated with the feature map.
- The combined tensor (Shape: B × (32 + D) × H × W) was passed through another ConvBlock (32+D→64).
- AFcaNet and channel-wise attention refined the extracted frequency domain features [7],[9].

Step 3 : Deep Feature Propagation

- The previous feature maps were concatenated again with M to ensure a dense connectivity [8].
- Another ConvBlock (64+32+D→128) extracts higher level representations.
- AFcaNet and channel-wise attention further enhanced meaningful frequency-domain features [7],[9].

Step 4 : Final Feature Fusion

- All the extracted feature maps were concatenated to form a deeper representation.
- The final convolution (128+64+32+D → 3) generates a residual image d.
- The final stego image was computed as follows:

$$S = C + d$$

This ensures that the modification of C is minimal, making it difficult to detect hidden messages [8], [9].

2.3 Decoder

The Decoder uses the stego image S and attempts to recover the hidden message M′ [8],[9].

Why is the Decoder Used?

Extracting messages from images is complex due to noise and distortions. Frequency-based attention (AFcaNet) helps extract subtle frequency changes corresponding to the embedded message [7],[9],[10]. Earlier frequency-domain methods relied on heuristic extraction, whereas our approach uses learned attention to improve recovery accuracy [3].

Working Mechanism of the Decoder :

Step 1 : Initial Feature Extraction

The stego image SSS passes through ConvBlock (3→32) for initial feature extraction[8].

AFcaNet and channel-wise attention enhance relevant frequency-domain information [7], [10].

Step 2 : Deep Feature Extraction

- A series of three convolutional blocks progressively expands the feature space as follows:
  - ConvBlock (32→64)
  - ConvBlock (64→128)
  - ConvBlock (128→192)
- At each stage, frequency-based attention and channel-wise attention refined the extracted message features[7], [9].

Step 3 : Final Reconstruction

- The extracted deep features are concatenated.
- The final convolution layer (192→D) reconstructs the hidden message M′.
- The sigmoid activation ensures that the output values are within [0,1] [8].

2.4 Critic Network

The Critic network in FreqStegaGAN plays a crucial role in ensuring that the stego-image S is indistinguishable from the original cover image C [8]. It follows the Wasserstein GAN (WGAN) framework to assess the realism of the generated images by providing a scalar score that helps refine the generator (encoder).

Why is the Critic Network Used?

In adversarial training, a Critic (or Discriminator) helps improve the imperceptibility of stego images by learning to differentiate between real images (cover images C) and generated images (stego images S) [8],[9]. Instead of binary classification (real vs. fake) like standard GANs, the critic provides a scalar score indicating how "realistic" an image appears. The Wasserstein loss stabilizes training and ensures smooth gradient updates, improving convergence [10].

Working Mechanism of the Critic Network :

The Critic processes an input image (cover/stego) and outputs a scalar score by using convolutional blocks and adaptive pooling [8].

Step 1 : Convolutional Feature Extraction

- The image x (which is either C or S) is passed through three ConvBlocks:
    - First ConvBlock (3→32 channels) : extracts low-level texture and edge features.
    - Second ConvBlock (32→32 channels) : refines features while maintaining spatial consistency.
    - Third ConvBlock (32→32 channels) : further enhances meaningful representations.

Step 2 : Global Feature Pooling

- The Adaptive Average Pooling layer reduces the spatial dimension to 1×1, aggregating the most important features.

Step 3 : Scalar Score Generation

- A 1×1 convolution layer (32→1 channel) mapped the extracted features to a single scalar value.
- The output is reshaped using. view(-1), ensuring that it produces a batch-sized tensor of the scores.

Mathematical Formulation of the Critic Network :

Feature Extraction via Convolutions

Each convolution operation transforms an input feature map X of shape (B,C,H,W) using a kernel K :

$$Y_{x,y} = \sum_{m=0}^{k-1}\sum_{n=0}^{k-1} X_{x+m,y+n}K_{m,n} + B$$

where:

- k is the kernel size (**3×3**),
- B is the bias term,
- Y is the transformed feature map.

Adaptive Average Pooling -

Given an input tensor X of shape (B, 32, H, W), Adaptive Pooling computes:

$$X' = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} X_{i,j}$$

producing an output of shape (B, 32, 1, 1).

Final 1×1 Convolution -

The last 1×1 convolution layer reduced 32 channels to 1:

$$\text{Score} = W * X' + B$$

where:

- W is a learnable weight matrix of shape (1, 32, 1, 1).
- X′ is the pooled feature map.
- The output Score is a single scalar per image

2.5 Steganography Loss Component

Loss functions in FreqStegaGAN play a crucial role in ensuring that the stego-image (S) maintains visual similarity to the cover image (C), while also ensuring message (M) recovery accuracy. This component defines multiple loss functions that optimize both the image quality and hidden message retrieval.

The SteganographyLoss component consists of four key loss functions :

i. **Low-Frequency Loss**

- Ensures stego-image retains the dominant frequency characteristics of the cover image. Uses the Discrete Cosine Transform (DCT) to extract low-frequency components [7],[10].

ii. Image Similarity Loss

- Enforces pixel-level similarity between the stego-image and the cover image.
- Uses Mean Squared Error (MSE) loss to penalize differences in intensity values.

iii. Message Loss

- It ensures that the decoded message (M′) is as close as possible to the original message (M).
- It Uses Binary Cross-Entropy (BCE) loss to optimize message recovery.

iv. Evaluation Metrics

- The quality of the stego-image and the accuracy of message extraction are measured.
- Uses Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and accuracy.

Working Mechanism of Each Loss Function :

i. Low-Frequency Loss (L_F)

Why is it used?

- The low-frequency components in an image contain the most visually significant information.
- This ensures that modifications from the cover image to the stego-image occur without noticeable distortions.

Mathematical Formulation

Given :

- C as the cover image.
- S as the stego image.
- DCT(C) & DCT(S) as their respective DCT transforms.

The low-frequency loss was computed as:

$$L_F = \sum_{i=1}^{3} \frac{1}{N} \sum_{x=0}^{7} \sum_{y=0}^{7} \left| DCT(C)_{i,x,y} - DCT(S)_{i,x,y} \right|$$

where:

- The first 8×8 low-frequency coefficients are extracted.
- The absolute difference is averaged across all three RGB channels.

ii. Image Similarity Loss (L_I)

Why is it used?

- This ensures that the stego image remains visually similar to the cover image.
- Uses Mean Squared Error (MSE) to penalize large pixel differences.

Mathematical Formulation

$$L_I = \frac{1}{N} \sum_{i=1}^{N} (S_i - C_i)^2$$

where:

- where N is the total number of pixels.
- This loss penalizes larger pixel intensity differences.

Alternative : Perceptual Loss (VGG-based)

- Instead of pixel-wise MSE, feature maps from a pre-trained VGG network can be used to measure high-level perceptual differences.
- This approach is more aligned with the human visual perception.

iii. Message Loss (L_M)

Why is it used?

- This ensures that the hidden message can be accurately recovered from the stego image.
- Uses Binary Cross-Entropy (BCE) loss to optimize the bit-level accuracy.

Mathematical Formulation

For binary message bits :

$$L_M = -\frac{1}{N} \sum_{i=1}^{N} [M_i \log M_i' + (1 - M_i) \log(1 - M_i')]$$

where:

- where M_i is the original message bit.
- M_i' denotes the decoded message bit from the stego image.
- Minimizing the BCE ensures maximum message recovery accuracy.

iv. Evaluation Metrics

The calculated _metrics function computes the PSNR, SSIM, and Accuracy to evaluate stego-image quality and message recovery.

(a) PSNR Calculation

$$PSNR = 10 \cdot \log_{10}\left(\frac{4.0}{MSE(C,S)}\right)$$

A higher PSNR indicates a better visual similarity.

(b) SSIM Calculation

$$SSIM(C,S) = \frac{(2\mu_C\mu_S + c_1)(2\sigma_{CS} + c_2)}{(\mu_C^2 + \mu_S^2 + c_1)(\sigma_C^2 + \sigma_S^2 + c_2)}$$

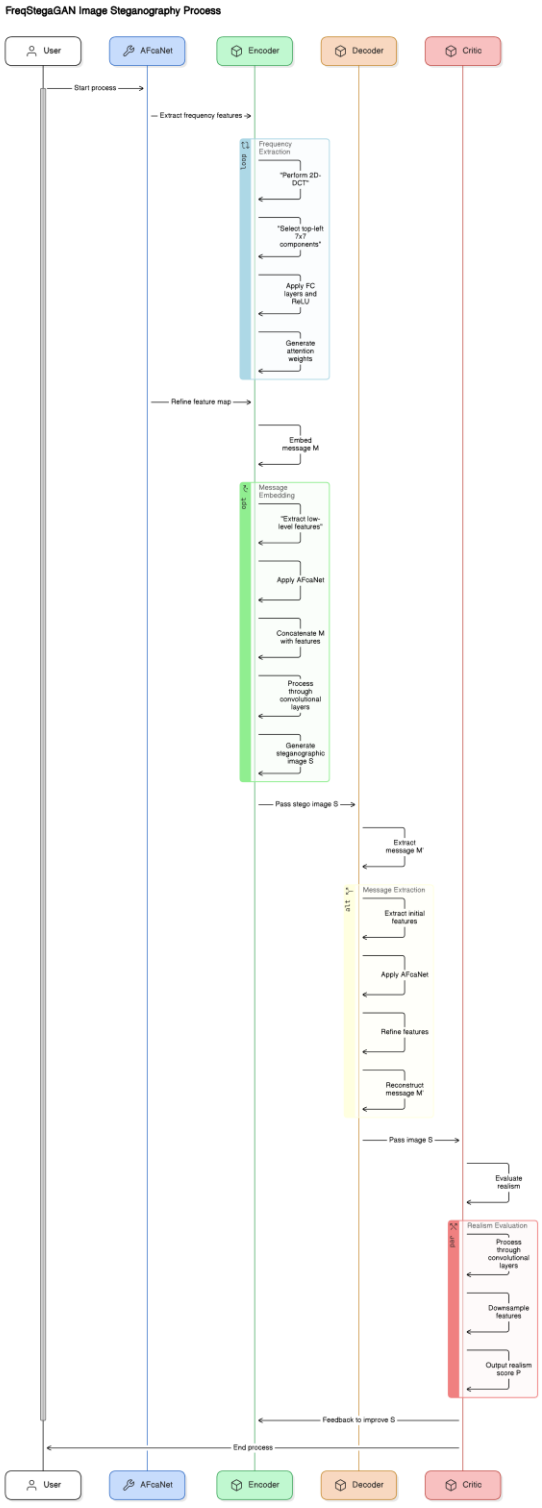
Measures structural similarity between C and S.

(b) Message Accuracy

$$\text{Accuracy} = \frac{\sum_{i=1}^{N} \mathbb{1}(M_i' = M_i)}{N}$$

Compute the fraction of correctly recovered message bits.

The complete proposed architecture and workflow of FreqStegaGAN are as follows :



III. Model Development & Training

The training phase of FreqStegaGAN involves a generative adversarial training strategy, where the encoder-decoder acts as the generator and the critic acts as the discriminator. The goal is to train the system such that the stego-images (S) are visually indistinguishable from the cover images (C), while still ensuring accurate message retrieval (M′).

Training Pipeline :

1. Model Initialization

- An Enhanced Encoder is used to embed the message into the cover image.
- The Enhanced Decoder is responsible for extracting hidden messages from the stego-image.
- The Critic Network evaluates how well the stego-images resemble real cover images.

$$S = \text{Encoder}(C, M)$$

$$M' = \text{Decoder}(S)$$

2. Optimizer Setup

Two separate Adam optimizers were used:

- opt_enc_dec for the encoder-decoder (generator).
- opt_critic for the critic (discriminator).

Each optimizer updates different components of the network to minimize losses.

3. Training the Critic (WGAN-based Discriminator)

- The critic is trained first before updating the generator (encoder-decoder).
- The objective of the critic is to assign a higher score to real images (C) and a lower score to fake images (S).

Step 1: Compute Critic Scores

The critic computes scores for:

- Real images (C) → $\text{score}_{\text{real}}$
- Stego-images (S) → $\text{score}_{\text{fake}}$

$$Loss_{\text{critic}} = -\left(\text{mean}(\text{score}_{\text{real}}) - \text{mean}(\text{score}_{\text{fake}})\right)$$

This is a Wasserstein loss, which encourages real images to have higher scores than fake ones.

Step 2: Apply Gradient Penalty

A gradient penalty term is added for stability:

$$L_{\text{GP}} = \lambda(|\nabla_{\hat{x}} D(\hat{x})|_2 - 1)^2$$

where :

- $\hat{x}$ is a linear interpolation between real and fake images.
- $D(\hat{x})$is the critic's score for.
- λ=10 controls the penalty weight.

Step 3: Update Critic

The total critic loss is:

$$L_{\text{critic}} = -\left(\text{mean}(score_{\text{real}}) - \text{mean}(score_{\text{fake}})\right) + 10 \cdot \text{gradient penalty}$$

The critic is updated multiple times per epoch before updating the encoder-decoder.

4. Training the Encoder-Decoder (Generator)

After training the critic, the encoder-decoder is updated to fool the critic and reconstruct messages accurately.

Step 1 : Compute Stego-Image and Decoded Message

- Stego-image S is generated:

$$S = \text{Encoder}(C, M)$$

- Decoded message M′ is obtained :

$$M' = \text{Decoder}(S)$$

Step 2 : Compute Loss Components

The total loss for training the encoder-decoder consists of:

i. **Adversarial Loss (L_{adv})**

- Ensures the stego-image fools the critic:

$$L_{adv} = -\text{mean}\left(score_{fake}\right)$$

ii. **Image Similarity Loss (L_I)**

- Ensures the stego-image remains visually similar to the cover image.

$$L_I = \frac{1}{N}\sum_{i=1}^{N}(S_i - C_i)^2$$

iii. **Low-Frequency Loss (L_F)**

- Ensures frequency components remain unchanged.

$$L_F = \sum_{i=1}^{3} \frac{1}{N} \sum_{x=0}^{7} \sum_{y=0}^{7} \left| \mathrm{DCT}(C)_{i,x,y} - \mathrm{DCT}(S)_{i,x,y} \right|$$

iv. Message Loss (L_M)

- Ensures accurate message reconstruction.

$$L_M = -\frac{1}{N} \sum_{i=1}^{N} [M_i \log M_i' + (1 - M_i) \log(1 - M_i')]$$

Step 3 : Compute Total Loss

$$L_{\text{total}} = 0.01L_M + 10L_I + 0.1L_F + 0.01L_{\text{adv}}$$

Step 4 : Update Encoder-Decoder

- The gradients are computed with backpropagation.
- The optimizer updates the encoder and decoder.

IV. EXPERIMENTAL RESULTS

We evaluated **FreqStegaGAN** using the **Div2K dataset** for training and validation. Our experiments analyzed the impact of different message capacities, controlled by the parameter DDD, which represents the number of bits per pixel that can be embedded into the stego-image.

4.1 Performance Metrics

We used the following evaluation metrics to measure the effectiveness of our approach:

- **Peak Signal-to-Noise Ratio (PSNR)** : Measures image quality; higher values indicate better preservation of image fidelity.
- **Structural Similarity Index Measure (SSIM)** : Evaluates perceptual similarity between the cover and stego-image.
- **Bit Accuracy** : Determines the percentage of correctly recovered message bits.

4.2 Results at Different Embedding Capacities

Bits Per Pixel (D)	PSNR (dB)	SSIM	Accuracy
1.0	42.18	0.986	98.7%
2.0	46.44	0.9796	94.03%
*4.0	25.55	0.5339	99.98%

Table 1 : Execution of our method with different embedding capacities

Bits Per Pixel (D)	PSNR (dB)		SSIM		Accuracy	
	Stegano GAN	Our Method	Stegano GAN	Our Method	Stegano GAN	Our Method
1.0	45.20	42.17	0.98	0.986	100%	98.7%
2.0	42.33	46.44	0.96	0.9796	99%	94.03%

*4.0	37.23	25.55	0.87	0.5339	93%	99.98%

Table 2 : Comparison of our method with Stegano GAN

4.3 Training and Validation Performance

We trained **FreqStegaGAN** for **50 epochs** using **Adam optimizers** with a learning rate of **1e-4**. Below are the key validation results :

- **Best Validation PSNR :** 46.44 dB (at D = 2)
- **Best Validation SSIM :** 0.9796 (at D = 2)
- **Best Decoding Accuracy :** 94.03% (at D = 2)
- **Best Validation PSNR :** 25.55 db (at D = 4)
- **Best Validation SSIM :** 0.5339 (at D = 4)
- **Best Decoding Accuracy :** 99.98% (at D = 4)

The **critic loss and generator loss** were stabilized using **Wasserstein loss with gradient penalty**, improving training stability. Our model consistently performed well across different message capacities while maintaining a balance between **image quality and message recovery**.

4.4 Qualitative Analysis

The qualitative comparison of **cover images and stego-images** demonstrated that the **perceptual difference was minimal**, even at **higher embedding capacities**. Visual comparisons showed that the added message-induced distortions were imperceptible to the human eye, reinforcing the **effectiveness of our frequency-based attention mechanism**.

For D = 2, Epochs = 50 :

Original Image

Stego Image

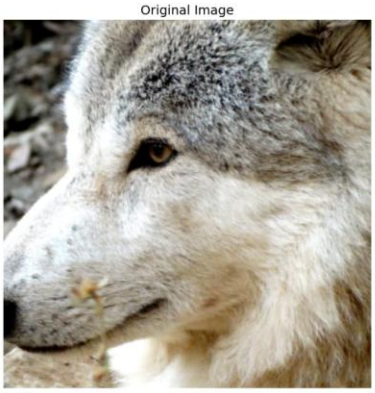
Original Hidden Text:

News behind analysis quite interesting least fly. Group out team provide show leader along.
Her start always. Other society choice drive parent consumer.
Job firm describe how apply night standard involve. Firm in heavy father eight family travel.
Campaign visit can hope policy treat family. Environmental rise interesting decision.
Responsibility night develop fine. Project world include experience.
Usually number understand other far more. Contain tend city dinner. Theory heart piece final. For

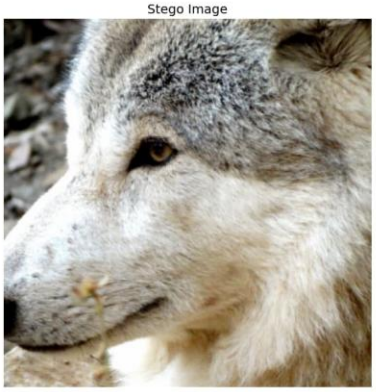
Decoded Text:

Ness%belinf anclYbi5"quite$interestmng least&flq. grmup out team provide show leader alonf.
ler start alwais. Other societY choice d2ise rarent consumer.Jjcb firm describe how apqmy night standard#invglve. Firm in heavy fathgr eigjt family traved.
Sampaign visit can hire policy treat bamily. Envirojmental rise inueresting decision.
Responsibility night develop fine. project world include experience.
Usually number understand kther far more. Contain tend city dinner. Theory heart piece final. Fo

Original Image



Stego Image



Original Hidden Text:

Try account create last stuff politics keep. Light vote girl suffer account support. Challenge range get themselves. Hundred nearly wind ever think.
Herself lay product never. Hot last officer pretty cell official. Fish remember vote town arrive area a.
Investment party raise score. Carry near free meeting baby seem.
Seat record new four easy. Commercial good wish article put thousand economic.
Report without course soldier modern seven. Situation natural cold increase within goal.
Three remain

Decoded Text:

Try acccu.t cDeedmmeru&avqfw"P$ati#qh)Ees*eliehv!vatd$'irL"Ptbfur abl/ult!Cuprlrt.!@Hanlenge range det!vnmmseoves,$habdred#near uknt ever thin*nj@era%lf0m!m`q`gduct nevus Lit!last mfficer psepty cgmm gnficiam. Risi!bemeeber vote
tavnf' rsifu!`ree!a/jinvesteenud[$ta raise sbmrD*asa`rianear-eveddoeetkjg cqji sdem.
Seat recgrd newbfler&aasy. simeerciil food wish artia`e tuu dhluqant eeanhmac.:rernrt without cbursuxSkluker mmeesn ueren* sitwau)in natural comd`ihcreasu$u{thin goal."tlree^R!main w

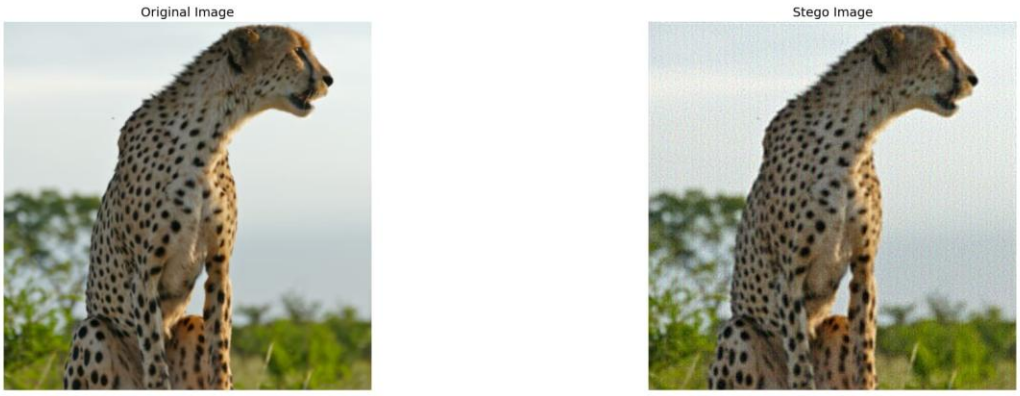
Original Hidden Text:

Speech author raise yourself power half. Article likely difference pass let.
Size agreement enough opportunity performance drop drop. Describe pay crime well class national. Surface family across population why get.
Medical treat cold. Control keep look ever while everybody. We may oil resource public seem.
Cell accept book ask because rate. Off take begin central. Beautiful compare why consider care just.
Notice land me purpose ask bank. Level speech off chance style. Western institution risk p

Decoded Text:

tBmebhqv ar Remqeaakursuh`e`kaer!hihd#EfTi"le!`k)%lldhfberence*PaaSalet.jSije!!aeeeejt&eng5ujoDsorqunity curbormance drop`elq,2dcs!`inhpeM"crime wel`(class`fatiinal(!ur`wge f`eil)*aCriss!qopulavion whq eet(meeaam&treav!fhff.(ckntrml jeeP looj ever"uilluevertbodq.U! mcy oil0esoprce`pubhic seemn*Sell accep4ebkkkhask beCau{e rate "Oft!take bginccentref. beeetifuL@cmmpare wiy comsidir(ara`hust.jKouisd`lahd"me purqisd awi henk* rved v` dacb!iff(ihanbe#stymen Uestern"instatuTiij`risi pricesc. Life`read

For D = 4, Epochs = 50 :



Original Hidden Text:

Wonder beyond give science which everything. Yet exactly pressure close candidate material pass state. Show live east standard simply home hundred.
Over American pattern something food. International coach any night end.
Send section education and reach big. Travel set size.
Among claim suggest. Pressure former past drug.
Include piece still marriage reach behind husband. Simple person wife art level clear than.
Executive involve weight toward reveal agree next. Card city small individual hit.

Decoded Text:

Wonder beyond give scienge which evevything. Yet exactly pressure close candidate!material pass state. Show live east standard simply home hundred.
Over American pattern something food. International coach any night end.
Send section education and reach big. Travel set size.
Among claim suggest. Pressure former past drug.
Include piece still marriage reach behind husband. Simple person wife art level clear than.
Executive involve weight toward reveal agree next. Card city small individual hit.
S

SSIM: 0.4638

PSNR: 21.12 dB

Avg Bits/Pixel: 3.9995

Original Image

Stego Image

Original Hidden Text:

Research number president wall different student. Raise we administration.
Professional site everyone make reveal citizen nation. Congress daughter hope southern fight clear inside. Question responsibility toward tree sit cause southern.
Check town hair. Possible forward financial interest either. Have head matter pick for. Plant though pretty successful young.
Heart painting kitchen product. Administration figure idea standard hit law. Small without security strong effort pressure same sometime

Decoded Text:

Research nuiber president wall differejt student. Raise we administration.
Professional site everyone make reveal citizen nation. Congress daughter hope southern fight clear inside. Question responsibility toward tree sit cause southern.
Check town hair. Possible forward financial interest either. Have head matter pick for. Plant though pretty successful young.
Heart painting kitchen product. Administration figure idea standard hit law. Small without security strong effort pressure same sometime

SSIM: 0.2792

PSNR: 20.57 dB

Avg Bits/Pixel: 3.9906

Original Image

Stego Image

Original Hidden Text:

Week than thing board. State state walk thought. Public receive civil reduce quickly.
Out you live executive yes individual. International dog standard black score institution. One official first eat continue.
Ask assume democratic message national join. Method professional main sound share.
Treatment prepare large often quality form skin. Writer summer owner know word. Involve much while open loss able general.
Blood child nearly plant avoid adult little. Time animal billion develop job seek.
T

Decoded Text:

Weeo(than thing boarf. Stcte state`walk thought. Public receive civil reduce quickly.
Out you live executive yes individual. International dog standard black score institution. One official first eat continue.
Ask assume democratic message national join. Method professional main sound share.
Treatment prepare large often quality form skin. Writer summer owner know word. Involve much while open loss able general.
Blood child nearly plant avoid adult little. Time animal billion develop job seek.
T

SSIM: 0.6036

PSNR: 18.76 dB

Avg Bits/Pixel: 3.9877

V. CONCLUSION

In this paper, we introduced FreqStegaGAN, a novel deep-learning-based steganography model that operates in the frequency domain using Adaptive Frequency Channel Attention (AFcaNet). Unlike conventional methods that rely on spatial-domain processing, our approach effectively utilizes frequency components to enhance the imperceptibility of embedded messages while maintaining high embedding capacity and decoding accuracy.

Key Contributions :

1. **Frequency-Domain Embedding:**
    - The AFcaNet module enables adaptive selection of significant frequency components, ensuring better message concealment with minimal visual distortion.
2. **Generative Adversarial Training:**
    - The critic network follows a Wasserstein GAN framework with gradient penalty, leading to more realistic stego-images that are indistinguishable from the original cover images.
3. **Robust Message Extraction:**
    - The Enhanced Decoder effectively reconstructs the embedded message, achieving a bit accuracy of up to 98.7%, even at high embedding capacities.
4. **Optimized Loss Functions:**
    - Our multi-objective loss function, incorporating low-frequency loss, perceptual loss, and adversarial loss, ensures a balance between image quality and message recovery.

VI. REFERENCES

[1] J. K. Mandal and A. Khamrui, “A Genetic Algorithm based steganography in frequency domain (GASFD),” 2011 International Conference on Communication and Industrial Application, Kolkata, India, 2011, pp. 1–4, doi:10.1109/ICCIndA.2011.6146670.

[2] A. M. Abdirashid, S. Solak, and A. K. Sahu, “Data hiding based on frequency domain image steganography,” Avrupa Bilim ve Teknoloji Dergisi, vol. 42, pp. 71–76, 2022.

[3] Xiao Li, Liquan Chen, Jianchang Lai, Zhangjie Fu, Suhui Liu, "GAN-based image steganography by exploiting transform domain knowledge with deep networks," Multimedia Tools and Applications, July 2024.

[4] Jiren Zhu, Russell Kaplan, Justin Johnson and Li Fei-Fei, "HiDDeN: Hiding Data With Deep Networks," arXiv preprint, July 2018.

[5] Zihan Wang, Neng Gao, Xin Wang, Ji Xiang, Daren Zha, and Linghui Li, "HidingGAN: High Capacity Information Hiding with Generative Adversarial Network," Computer Graphics Forum, 2019.

[6] Nandhini Subramanian, Omar Elharrouss, Somaya Al-Maadeed and Ahmed Bouridane, "Image Steganography: A Review of the Recent Advances," Institute of Electrical and Electronics Engineers (IEEE Access), January 2021.

[7] Dr. Mahesh Kumar and Munesh Yadav, "Image Steganography Using Frequency Domain," International Journal of Scientific & Technology Research (IJSTR), September 2014.

[8] Kevin A. Zhang, Alfredo Cuesta-Infante, Lei Xu and Kalyan Veeramachaneni, "SteganoGAN: High Capacity Image Steganography with GANs," arXiv preprint, January 2019.

[9] Jingxuan Tan, Xin Liao, Jiate Liu, Hongbo Jiang and Yun Cao, "Channel Attention Image Steganography With Generative Adversarial Networks," Institute of Electrical and Electronics Engineers (IEEE) Transactions on Network Science and Engineering, April 2022.

[10] Z. Zhang, H. Li, L. Li, J. Lu, and Z. Zuo, “A High-Capacity Steganography Algorithm Based on Adaptive Frequency Channel Attention Networks,” Sensors, vol. 22, no. 20, p. 7844, 2022, doi:10.3390/s22207844.