

Article

Not peer-reviewed version

---

# Content Protection Method for Remote Sensing Images Based on Deep Information Hiding

---

[Peng Luo](#)\*, Jia Liu, Jun de Mu

Posted Date: 12 January 2024

doi: 10.20944/preprints202401.0925.v1

Keywords: remote sensing image; distribution; deep information hiding; attention mechanism.






Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Article

# Remote Sensing Image Secure Distribution Scheme Based on Deep Information Hiding

Peng Luo <sup>1,2,\*</sup> , Jia Liu <sup>2</sup>  and Dejun Mu <sup>1</sup> 

<sup>1</sup> School of Cybersecurity, Northwestern Polytechnical University, Xi'an, China

<sup>2</sup> School of Cryptography Engineering, Engineering University of PAP, Xi'an, China

\* Correspondence: lp\_nwpu@mail.nwpu.edu.cn

**Abstract:** To ensure the security of highly sensitive remote sensing images (RSI) during their distribution, it is essential to implement effective content security protection methods. Generally, secure distribution schemes for remote sensing images often employ cryptographic techniques. However, sending encrypted data exposes communication behavior, which poses significant security risks to the distribution of remote sensing images. Therefore, this paper introduces deep information hiding to achieve secure distribution of remote sensing images, which can serve as an effective alternative in certain specific scenarios. Specifically, the Deep Information Hiding for RSI Distribution (hereinafter referred to as DIH4RSID) based on encoder-decoder network architecture with Parallel Attention Mechanisms (PAM) by adversarial training was proposed. Our model is constructed with four main components: a preprocessing network (PN), an embedding network (EN), a revealing network (RN) and a discriminating network (DN). The PN module is primarily based on Inception to capture more details of RSI and targets of different scales. The PAM module obtains features in two spatial directions to realize feature enhancement and context information integration. The experimental results indicate that our proposed algorithm achieves relatively higher visual quality and secure level compared to related methods. Additionally, after extracting the concealed content from hidden images, the average classification accuracy is unaffected.

**Keywords:** remote sensing image; distribution; deep information hiding; attention mechanism

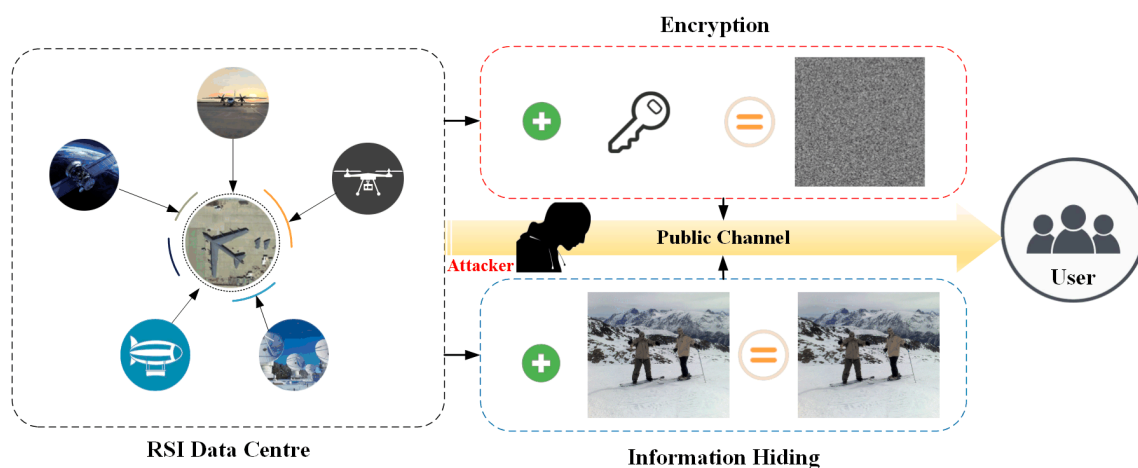
## 1. Introduction

As advancements in spatial, informational, and communication technologies persist, the primary method for applying spatial information has shifted towards the utilization of digital products and network integration. [1]. This has greatly facilitated the communication and sharing of remote sensing information, allowing for further utilization of remote sensing images. However, it has also introduced security concerns during the distribution of such information through networks [2]. RSI obtain data from vulnerable sites like military installations, oil fields, and airports, posing a risk for potential misuse and theft. Moreover, processing these images without implementing proper security measures can simplify unauthorized data retrieval and offer uninterrupted access to confidential information, which could subsequently be used for illicit purposes [3]. To safeguard the confidentiality of highly sensitive remote sensing data, adopting a robust content security protection strategy is essential. Encryption is commonly acknowledged as a prevalent technique for protecting the confidentiality of remote sensing images during their transmission and storage. [4]. That is the use of symmetric or public key cryptographic algorithms to transform plaintext into ciphertext, which is then transmitted through a public channel to the recipient. Because ciphertext presents itself as unintelligible or scrambled text, it readily draws the focus of attackers monitoring the communication channel. This can lead to two consequences: firstly, if the attacker intercepts the ciphertext, they can attempt to decrypt it using various attack methods against different cryptographic systems [5].

Secondly, if the attacker is unable to decrypt the ciphertext, they may disrupt the channel to prevent distribution. Indeed, it is necessary to explore alternative methods for secure distribution of remote sensing images in certain specific cases. The information hiding technology in the field

of information security has begun to attract more and more attention [6,7]. Unlike encryption, in information hiding scheme, secret information is embedded in seemingly harmless host information, and attackers cannot intuitively determine whether the information they are monitoring contains secret information. In other words, hosts containing hidden information will not attract attacker's attention and suspicion. The purpose of information hiding is to make enemies unaware of where there are secrets, as it hides the form of information that exists. The comparison between the idea of our scheme and the traditional encryption distribution method is shown in Figure 1. Although there are relatively few publicly published literature on the use of information hiding technology for remote sensing image distribution, some scholars' research results [8–12] provide us with good insights. These methods mainly achieve the hiding of one image within another of the same size; however, they do not specifically address discussions and research regarding remote sensing images. Therefore, this paper will draw on their ideas to design a new algorithm suitable for the secure distribution task of remote sensing images. The primary contributions of our proposed scheme are as follows:

- To our knowledge, DIH4RSID is the first to explicitly propose the use of deep information hiding technology to ensure the secure distribution of RSI. Therefore, our method not only opens a new way of thinking about RSI security distribution, but also enriches the research results in this field;
- Unlike the existing HIWI(Hiding images within images) framework, our study proposes a novel preprocessing network architecture, which is designed based on Inception networks and crafted to conform to the unique properties of remote sensing images, which can capture detailed information of objects at different scales.
- According to the characteristics of our tasks and remote sensing images, a new attention mechanism PAM is designed in this paper, which carries out two kinds of pooling from two dimensions respectively. Convolution operations can then capture cross-channel relationships and spatial remote dependencies.
- A discriminator is added to the scheme, and iterative training is carried out by WGAN-GP, which improves the stability and correct convergence speed of the model.



**Figure 1.** The comparison between our scheme and the traditional encryption distribution method. The encrypted remote sensing image is presented in garbled code, which is easy for attackers to pay attention to. In our scheme, the remote sensing image is hidden in the ordinary image in an invisible form, which can be more effective and secure distribution. .

## 2. Related Work

The approach in this article draws on many concepts and methods from the field of image information hiding; therefore, our analysis of related work is confined to the domain of information hiding. The origins of information hiding technology can be traced back to ancient covert means

of communication, such as invisible ink and miniaturized fonts. With the development of time, information hiding has gradually changed from traditional physical means to digital technology. Based on the application context, information hiding can be divided into digital steganography and digital watermarking. Generally speaking, the former is mainly used for covert communication, while the latter is used for copyright protection. According to such categorization, our DIH4RSID is more closely related to the field of digital steganography. The research of steganography can be roughly divided into three stages. Early information hiding technologies were mainly based on non-adaptive hiding strategies, among which the most typical representative is LSB (Least Significant Bit) [13]. LSB is a steganography method that modifies and stores information based on the least significant bit of an image. Using the insensitivity of human eyes to color differences, the secret information is put into the least significant bit of the picture by a certain embedding method, so that the information we need to hide is put into the least significant bit of the picture by a certain method. Because non-adaptive steganography does not consider the characteristics of the cover image itself, it is not safe and easy to be detected and analyzed. Based on this, adaptive steganography came into being, representing the second stage of steganography. Adaptive steganography considers the properties of the cover image itself, such as texture information and edge information of the image content. According to the characteristics of difficult detection of complex areas of image texture, secret information is selectively embedded into areas with complex textures or rich edges of the cover, which improves the anti-steganographic detection ability of loaded images. At the same time, all kinds of adaptive steganography algorithms are combined with STC [14] encoding methods, the difference is that the distortion function is different. Such algorithms are represented by HUGO [15], WOW [16], UNIWARD [17] and HILL [18]. Although the adaptive steganography methods have achieved high performance, they are confronted with several challenges for both content adaptive based and statistics-based approaches. Firstly, Such algorithms can only embed a small number of bits or text information, and cannot embed multimedia information such as images [12]. At the same time, these methods often require specialized knowledge to design elaborate distortion cost functions. With the continuous development of analysis algorithms based on deep learning, the security of these traditional human-designed information hiding algorithms faces great challenges. This makes researchers begin to turn their attention to deep learning, attempting to use deep learning's powerful feature fusion ability to realize information hiding. Frameworks for information hiding based on deep learning, such as HiDDeN [19] and SteganoGAN [9], have been developed to accomplish the tasks of hiding and extracting information. This development signifies the progression of information hiding into its third phase, known as deep information hiding. These frameworks eliminate the need for manual design of embedding strategies and achieve higher payloads. However, they still only enable the covert transmission of small amounts of data. To address the challenge of hiding large image data, Baluja [20] presented a system to embed a full-color image into another of identical size while minimizing the quality degradation of both images. This is achieved by concurrently training deep neural networks to carry out both the embedding and extraction processes, which are specifically tailored to function in tandem. While this approach represents a significant innovation and yields impressive visual results, its robustness against analytical attacks leaves something to be desired. Rehman et al. [8] endeavored to develop an encoder-decoder architecture rooted in convolutional neural networks, accomplishing complete network training through the adoption of a novel loss function. While this approach proficiently conserved the fidelity of the concealed image, the visual quality of the crafted stego-image was subpar. In addition, in order to further improve the hiding performance, Duan et al. [21] introduced a reversible information concealing network that utilizes a U-Net architecture. The approach yielded pleasing outcomes in synthesizing concealed images as well as in the accurate retrieval of secret images. Nonetheless, their research did not delve into an in-depth examination of security concerns. Chen et al. [12] posited that certain secret images might possess intricate spatial characteristics. To address this, he proposed a multi-tiered robust auxiliary module aimed at augmenting the feature representation, subsequently elevating the restoration quality of secret

images. However, due to the absence of a discriminator within the framework, the enhancement in performance was not markedly evident. Some researchers introduce the attention mechanism into the field of deep information hiding [10,22–24], and promising results have been achieved. Tan et al. [23]. propose a new end-to-end image network architecture based on a channel-attention mechanism that generates adversarial networks. Steganography can produce perceptively indistinguishable steganographic images of different capacities. However, their programs cannot be used to directly embed and carrier such large remote sensing images. According to the above analysis, the existing method can not be directly applied to the safe distribution of remote sensing images, so it needs to be further modified to adapt to this task.

### 3. Proposed Scheme

In this section, the general structure of DIH4RSID is introduced first, and then, all parts of DIH4RSID are described in detail. Finally, we delineate the various loss functions and outline the training methods used.

#### 3.1. overview

As show in Figure 1, DIH4RSID workflow can be described as follows. Firstly, Alice(RSI owner) extracts feature maps through preparing network, and then hide that into a nature image (cover) by embedding network, which outputs a stego (also called the hidden image ). The stego could transmitted through public channels to Bob (receiver) , who decoded information through extracting network to reconstruct RSI. During the embedding process, discriminating network acts as a attacker to improve the indistinguish ability between cover and stego.

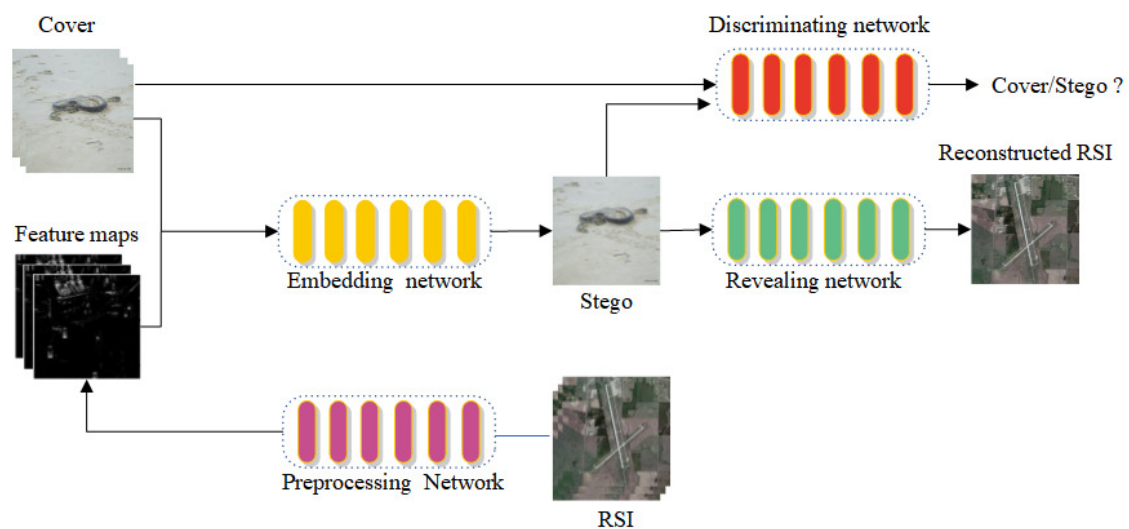


Figure 2. DIH4RSID Flowchart.

The secure distribution schema for RSI based on information hiding must satisfy there properties: 1) To avoid attracting the attention of attackers, embedding process should have minimal visual impact on cover, 2)In order to enhance the practicality of the algorithm, the semantics of the extracted remote sensing images should be preserved and 3) To improve the security of the model, it is important to restrict the success rate of detecting algorithms to a minimal level. Considering above mentioned three requirements, We have adopted three unique designs based on traditional encoder-decoder network, namely the preprocessing network based on Inception structure, PAM, and semantic loss.



### 3.2. Preprocessing network

RSI feature large scale variations in objects, rich details, complex structures, and ambiguous distinctions between subjects and backgrounds. The model proposed by Balujia and others employs a single-scale convolutional kernel to extract the spatial distribution features of image targets, which does not sufficiently delve frequency structures and spatial characteristics. To address this issue, we integrate the Inception structure into the design of the preprocessing network, as depicted in Figure 3. The Inception structure employs three distinct sizes of convolutional kernels and one max pooling in parallel to obtain spatial features at different scale receptive fields. In the parallel branches, dimensions are firstly adjusted through a  $1 \times 1$  convolution, then subjected to convolution kernels of varying scales and pooling operations. Finally, the extracted multi-scale spatial features are fused through Concat operations, enhancing the network's ability to capture a more complete description of target spatial structural features at different depth levels.

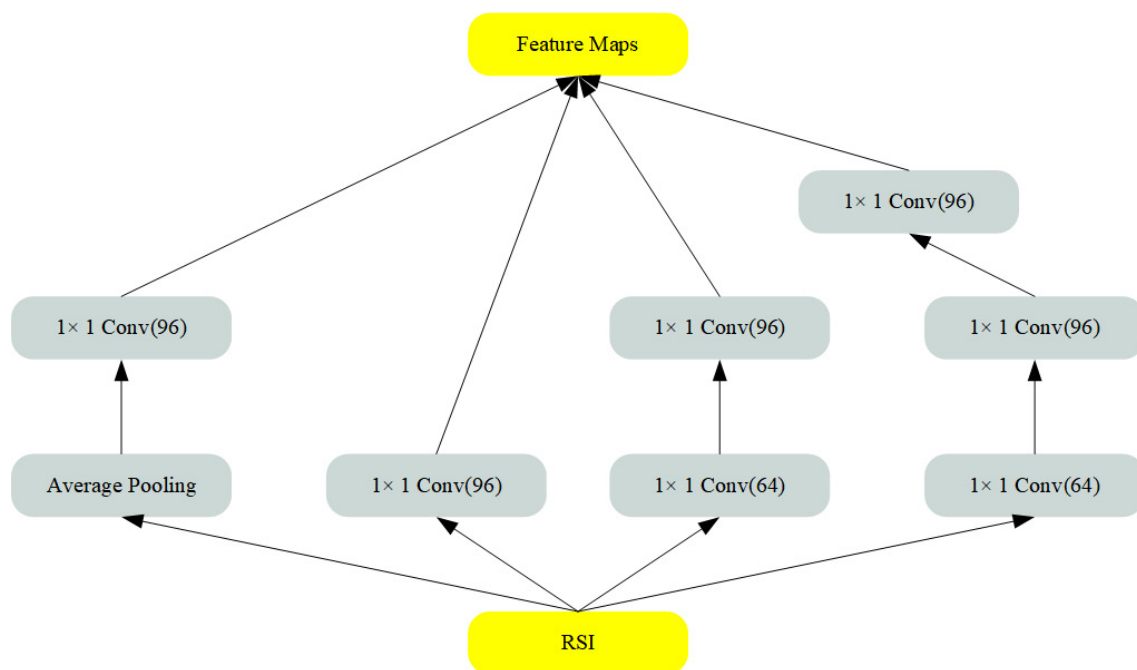


Figure 3. Preparing network architecture.

### 3.3. Parallel Attention Mechanism

In the realm of deep learning, the attention mechanism trains networks to concentrate on salient features while disregarding those that are irrelevant. For CNN-based information hiding schemes that directly produce stegoimages, the secret information is ingrained through integration with the cover features. Given that the value of these features varies with respect to information hiding, applying an attention mechanism could enhance the system's performance. Given that the perceptual signals of primary targets in RSI are typically concentrated in certain areas [25], spatial attention mechanisms are highly appropriate for our endeavor. They enable the model to concentrate on areas of the image that are most significant. However, calculating weights between features at all positions will bring significant computational effort. Consequently, drawing inspiration from the Coordinate Attention (CA) and Efficient Channel Attention (ECA) mechanisms, we design the PAM to effectively grasp the inter-channel relationships and extensive spatial dependencies, incorporating precise positional details. The detailed structure of PAM is shown in Figure 4.

Firstly, the horizontal average pooling, vertical average pooling, horizontal max pooling and vertical max pooling are performed for each channel of the input feature map denoted as  $F(h, w, c)$  respectively, as are formulated by equations (1)-(4).

$$F_{hap}(h, c) = \frac{1}{W} \sum_{i=1}^w F(h, i, c) \quad (1)$$

$$F_{vap}(w, c) = \frac{1}{h} \sum_{i=1}^h F(i, w, c) \quad (2)$$

$$F_{hmp}(h, c) = \max_{0 \leq i < w} F(h, i, c) \quad (3)$$

$$F_{vmp}(w, c) = \max_{0 \leq i < h} F(i, w, c) \quad (4)$$

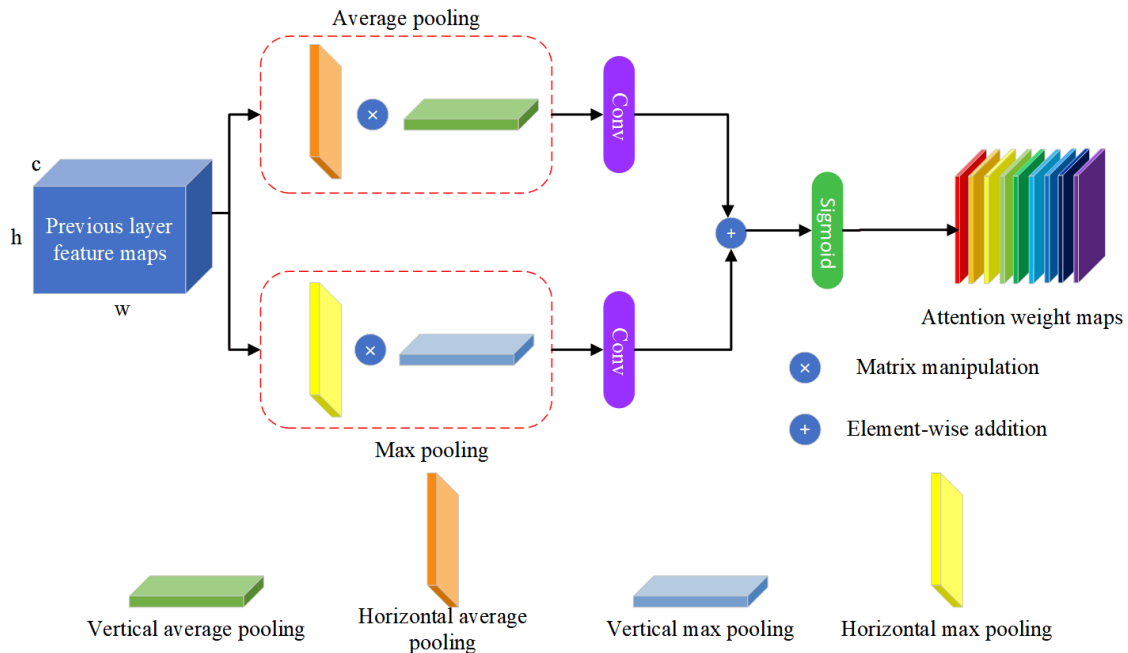
Subsequently, to prevent the reduction in channel dimensionality during cross-channel interactions, we employ Conv2D with a flexible kernel size. This is designed to produce attention weights across both spatial dimensions. Additionally, we incorporate two distinct types of pooling strategies. These processes collectively can be expressed as equation (5) - (6).

$$\hat{F}_{AP}(h, w, c) = \text{Conv2D}(F_{vap} \times F_{hap}, \text{kernal}) \quad (5)$$

$$\hat{F}_{MP}(h, w, c) = \text{Conv2D}(F_{hmp} \times F_{vmp}, \text{kernal}) \quad (6)$$

$$\text{kernal} = \left\lfloor \frac{\log_2(c)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (7)$$

According to ECA, the kernel of Conv2d is computed by (7), where  $\|_{\text{odd}}$  denotes as the nearest odd number,  $b$  and  $\gamma$  are set 1 and 2.



**Figure 4.** Detailed structure of PAM.

Finally, to determine the weight of the coordinate attention, we combine the adaptive pooling feature,  $\hat{F}_{AP}$  and  $\hat{F}_{MP}$  are combined using the operation of matrix multiplication. This procedure can be mathematically represented as formula (8).

$$Weight_{pam}(h, w, c) = \text{sigmoid}(\hat{F}_{AP} + \hat{F}_{MP}) \quad (8)$$

### 3.4. Embedding Network

We use a network structure similar to Dense Connection Architecture (DCA) as the backbone of the embedding network. There are several reasons why such architecture is suitable for DIH4RSID: (1) DCA reduces the problem of gradient disappearance by means of cross-layer connection, so that information will not be lost during transmission. This is important for DIH4RSID because remote sensing images typically have large sizes and complex backgrounds that require the network to be able to efficiently transmit and utilize feature information. (2) DCA can better utilize and enhance features in the transmission process by including the features of all previous layers in the input of each layer. This is also very beneficial for DIH4RSID, because the features of RSI are often complex, and the network needs to be able to extract and transmit these features effectively. (3) DCA combines features of previous layers to form richer description and discrimination of features. This enables the network to use the feature information more effectively, thus improving the hiding effect and extraction quality of DIH4RSID. (4) Reducing the number of parameters: DCA reduces the number of parameters in the model by reducing the number of connections, thus reducing the complexity and calculation cost of the model. This is also important for DIH4RSID because RSI is usually very large and requires low model complexity and computational cost to achieve effective classification.

To maintain stealthiness, it is crucial for the stego image to bear a close resemblance to the cover image. Moreover, to ensure the integrity of the recovered image, it should be nearly identical to the RSI. Nevertheless, the process involving various convolutional and activation layers can unavoidably lead to a loss of information from input images like cover images and RSI, which is detrimental to both hiding and revealing capabilities. To mitigate this issue, the introduction of both global and local skip connections is proposed, as are shown in Figure 5. The global skip connection facilitates the direct transmission of original image data to the uppermost layer, aiding in the enhancement of edge and texture detail synthesis, which in turn boosts both concealment and retrieval efficiency. On the other hand, the local skip connection allows for the unimpeded flow of RSI within the embedding network, ensuring that its details and semantic content are effectively incorporated into the stego. This incorporation aids the follow-up extraction network in precisely restoring the RSI. The embedding procedure is defined by a specific equation.

$$Stego = EN(C, PN(RSI, \theta_{PN}), \theta_{EN}) \quad (9)$$

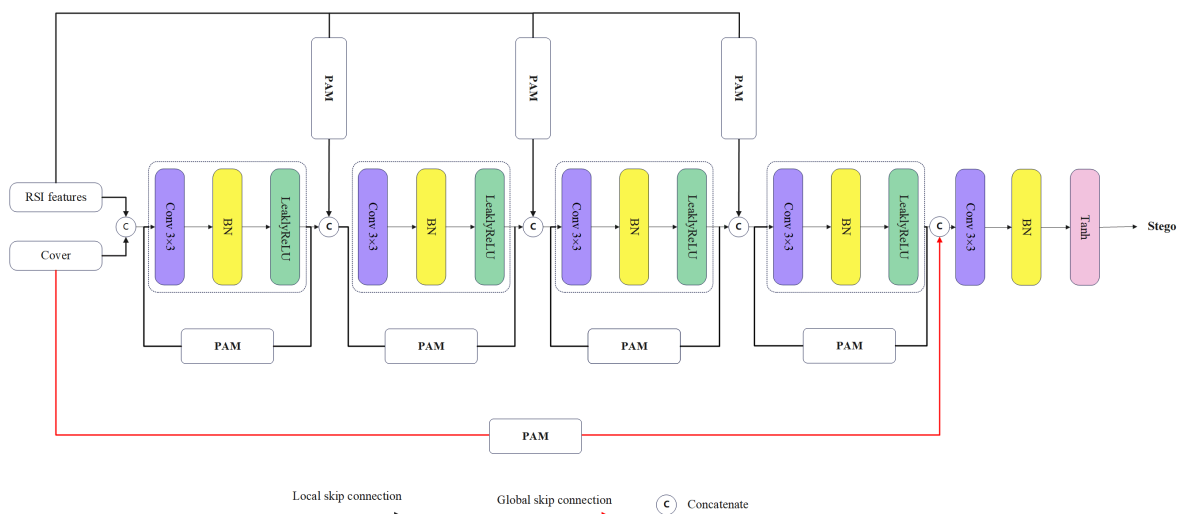


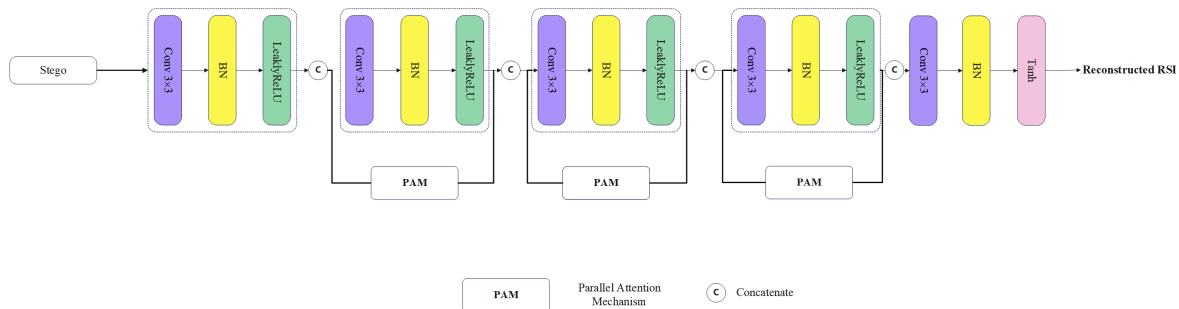
Figure 5. Embedding network architecture.



### 3.5. Revealing Network

RSI extraction is an inverse process of embedding. The revealing Network (RN) adopts a structure similar to the embedded network and still extracts depth features with dense connections, hoping to recover the RSI with high accuracy. Unlike the embedding process, the revealing process does not consider the effect of RSI feature integration, so the modules of global skip connection and local skip connection used in the embedded network are not used in the RN.

$$R\hat{S}I = RN(Stego, \theta_{RN}) \quad (10)$$



**Figure 6.** Detailed structure of revealing network.

### 3.6. Discriminating Network

The main purpose of this paper is to realize the safe distribution of remote sensing images, so not only to ensure the visual imperceptibility, but also to ensure that there are no obvious statistical distribution anomalies. To achieve this, we added an authentication network to the entire training process, which is described in Algorithm 1. In the selection of the discriminator network module, we mainly use the design idea of XuNet [26] network for reference, and make some changes. The first is to increase the number of nodes at the input side to handle 3D stego, and the second is to add ASPP [27] pooling technology which uses spatial pyramid pooling to obtain context information of different scales and residual information to improve the model's ability to perceive and distinguish stego. The detail of DN is listed in Table 1 and the process of DN can be formulated in equation 11.

$$\text{Probabilities of classes} = DN(Stego, \theta_{RN}) \quad (11)$$

**Table 1.** Overall Architecture details of the Revealing Network.

Inputs	modules	Kernal	Outputs
Stego( $3 \times 256 \times 256$ )	HPF	$3 \times 5 \times 5$	Out1( $3 \times 256 \times 256$ )
Input1	Conv-ABS-BN-Tanh—Average	$3 \times 5 \times 5$	Out2( $8 \times 128 \times 128$ )
Input2	Conv-BN-Tanh—Average	$8 \times 5 \times 5$	Out3( $16 \times 64 \times 64$ )
Input3	Conv-BN-Tanh—Average	$16 \times 1 \times 1$	Out4( $32 \times 32 \times 32$ )
Input4	Conv-BN-Tanh—Average	$32 \times 1 \times 1$	Out5( $64 \times 16 \times 16$ )
Input5	Conv-BN-Tanh—Average	$64 \times 1 \times 1$	Out5( $128 \times 8 \times 8$ )
Input6	ASPP	$3 \times 3, 1 \times 1$	Out5( $2560 \times 1 \times 1$ )
Input7	Fully Connection	-	Out5( $2 \times 1 \times 1$ )
Input8	SoftMax	-	Probabilities of classes ( $2 \times 1 \times 1$ )

### 3.7. Loss Function Design

Our framework can be broadly seen as similar to a GAN structure, where the *PN*, *EN*, and *RN* sub-networks function as a cohesive unit working end-to-end in the pipeline, synchronously updating each other. Built upon the concept of mutual adversariality, *DN* is alternately optimized

against them. The loss of embedding and revealing process can be denoted as  $\mathcal{L}_{PN,EN,RN}$  which comprises of embedding loss  $\mathcal{L}_e$  and revealing loss  $\mathcal{L}_r$ , and adversarial loss  $\mathcal{L}_{gd}$ , while the loss of discriminating process can be denoted as  $\mathcal{L}_{DN}$  which consists solely of the adversarial loss. Given that both the generated *stego* and the recovered  $RSI'$  are fundamentally RGB images, the generation loss and recovery loss are composed of three parts: pixel-wise loss, structural loss, and perceptual loss. Each of these loss functions serves as a measurement between the two images, focusing on different aspects and complementing one another. Therefore, in our approach, we combine all three to enhance the embedding effect and the quality of recovery.

$$\mathcal{L}_e = \text{MSE}(\text{Cover}, \text{Stego}) + \text{SSIM}(\text{Cover}, \text{Stego}) + \text{MSE}(\text{VGG19}(\text{Cover}), \text{VGG19}(\text{Stego})). \quad (12)$$

$$\mathcal{L}_r = \text{MSE}(RSI, RSI') + \text{SSIM}(RSI, RSI') + \text{MSE}(\text{VGG19}(RSI), \text{VGG19}(RSI')). \quad (13)$$

In equation (12) and (13), MSE denotes Mean Squared Error, SSIM denotes Structural Similarity Index Measure and VGG19 represents a deep convolutional neural network architecture; In this context, VGG19 is utilized solely for the extraction of semantic features by employing its pre-trained model, without undergoing any additional training.

The  $DN$  determines the probability that the input stego belongs to the cover, and the larger the value is, the closer it is to the distribution of the cover. As a loss function, it is generally optimized by minimizing its value, so a negative sign is added before the probability value of the loss. Therefore, the adversarial loss as  $EN$  and  $PN$  can be expressed in accordance with equation(14).

$$\mathcal{L}_{gd} = -DN(\text{Stego}). \quad (14)$$

In short, the purpose of training  $PN, EN$  and  $RN$  is to optimize the loss function as expressed in equation(15).

$$\mathcal{L}_{PN,EN,RN} = \lambda_1 \mathcal{L}_e + \lambda_2 \mathcal{L}_r + \lambda_3 \mathcal{L}_{gd} \quad (15)$$

where  $\lambda_1, \lambda_2, \lambda_3$  are weight factors that adjust the proportion of different loss functions in the total loss function. The  $DN$  strives to reduce the predicted score for covers while increasing it for stegos. We optimize it by the loss function as expressed in (16):

$$\mathcal{L}_{DN} = (0 - DN(\text{Stego}))^2 + (1 - DN(\text{Cover}))^2. \quad (16)$$

### 3.8. Training Process

When training is complete and the model is deployed to a real-world application scenario, the stego will be disseminated through public channels. In theory, a completely secure distribution system requires stegos and cover to follow the same distribution. In fact, it is difficult to achieve exactly the same distribution in practical applications, generally using some way to measure the distance between the two distributions, the smaller the distance, the better the hiding effect. Therefore, minimizing the distance between the two distributions becomes our optimization goal for generating the stego. Under the guidance of this idea, the original GAN proposed by Goodfellow et al. [28] can optimize the generated distribution on the basis of the KL divergence. Subsequent researchers are dedicated to enhancing GAN by developing appropriate network architectures and introducing novel loss functions to mitigate its numerous shortcomings. Among them, Arjovsky et al. [29] discovered that the Wasserstein distance offers benefits over both  $KL$  and  $JS$  distances, leading to the introduction of Wasserstein-GAN (WGAN) to achieve more stable training processes. In the original version of WGAN, the 1-Lipschitz constraint was imposed via weight clipping, but this approach had several drawbacks, such as potentially leading to gradient vanishing or exploding, as well as limiting the capacity of the model. To overcome these limitations, WGAN-GP (Wasserstein GAN with Gradient Penalty) [30] was proposed. Building on the foundation of WGAN, WGAN-GP introduces a gradient penalty to

more effectively enforce the 1-Lipschitz constraint, resulting in several advantages: improved training stability, elimination of problems associated with weight clipping, enhanced sample quality during training, and simplified fine-tuning. As a result, we utilize WGAN-GP to more accurately align the generator's output distribution, which similarly facilitates the achievement of a stable training regimen. For our task, the Wasserstein-1 distance between the cover distribution  $P_{cover}$  and the stego distribution  $Q_{stego}$  is denoted by the following equation(17).

$$W(P_{cover}, Q_{stego}) = \inf_{\gamma \in \Pi(P_{cover}, Q_{stego})} \int_{\mathcal{X} \times \mathcal{Y}} \|x - y\| d\gamma(x, y). \quad (17)$$

---

**Algorithm 1:** Training DIH4RSID. We use default values of  $\lambda = 10, n_{DN} = 5, \alpha = 0.0001, \beta_1 = 0, \beta_2 = 0.9$ .

---

**Input:** Cover dataset  $\mathcal{X}$ , RSI dataset  $\mathcal{R}$ ; initial PN parameters  $w_{PN}$ , initial EN parameters  $w_{EN}$ , initial DN parameters  $w_{DN}$ , initial RN parameters  $w_{RN}$ ; batch size  $m$ , DN iterations per EN iteration  $n_{DN}$ ; the gradient penalty coefficient  $\lambda$ , the number of DN iterations per generator iteration  $n_{DN}$ , the batch size  $m$ , Adam hyperparameters  $\alpha, \beta_1, \beta_2$ , the total iterations  $N$

**Output:** Trained parameters  $w_{PN}, w_{EN}, w_{DN}, w_{RN}$

```

for  $t \leftarrow 0$  to  $N$  do
  for  $i \leftarrow 0$  to  $n_{DN}$  do
     $\mathcal{L}^{(i)} \leftarrow 0$ 
    for  $j \leftarrow 0$  to  $m$  do
      Sample  $x \in \mathcal{X}$ 
      Sample  $r \in \mathcal{R}$ 
      Sample  $\epsilon \in \mathcal{U}[0, 1]$ 
       $\bar{x} \leftarrow EN(Concat(PN(r), x))$ 
       $\hat{x} \leftarrow \epsilon \bar{x} + (1 - \epsilon)x$ 
       $\mathcal{L}^{(i)} \leftarrow \mathcal{L}_d(x, \hat{x}) + \lambda(\|\nabla_{\hat{x}} DN_{w_{DN}}(\hat{x})\|_2 - 1)^2$ 
    end
     $w_{DN} \leftarrow Adam\left(\nabla_{w_{DN}} \frac{1}{m} \mathcal{L}^{(i)}, w_{DN}, \alpha, \beta_1, \beta_2\right)$ 
  end
  Randomly select  $\{x_i\}_{i=1}^m$  a batch from  $\mathcal{X}$ 
  Randomly select  $\{r_i\}_{i=1}^m$  a batch from  $\mathcal{R}$ 
   $w_{PN} \leftarrow Adam(\nabla_{w_{PN}} \mathcal{L}_{PN, EN, RN}, w_{PN}, \alpha, \beta_1, \beta_2)$ 
   $w_{EN} \leftarrow Adam(\nabla_{w_{EN}} \mathcal{L}_{PN, EN, RN}, w_{EN}, \alpha, \beta_1, \beta_2)$ 
   $w_{RN} \leftarrow Adam(\nabla_{w_{RN}} \mathcal{L}_{PN, EN, RN}, w_{RN}, \alpha, \beta_1, \beta_2)$ 
end

```

---

In equation (17),  $W(P_{cover}, Q_{stego})$  represents the Wasserstein-1 distance between probability distributions of Cover and Stego.  $\Pi(P_{cover}, Q_{stego})$  represents the collection of all combined distributions  $\gamma$  that transport  $P_{cover}$  to  $Q_{stego}$ , where each  $\gamma$  must satisfy the marginal distributions to be consistent with  $P_{cover}$  and  $Q_{stego}$ .  $\|x - y\|$  represents the Euclidean distance between two points  $x \in Cover$  and  $y \in Stego$  in space, and the integral calculates the expected cost of moving from  $x$  to  $y$  over all possible transport plans  $\gamma$ . Based on the aforementioned principles, our training process is outlined in Algorithm 1. The (PN, EN, RN) and DN are trained alternately until the number of iterations reaches the maximum value., where (PN, EN, RN) jointly learn to minimize  $\mathcal{L}_{PN, EN, RN}$  while DN aims to minimize  $\mathcal{L}_{DN}$ . Note that DN is iterated 5 times once (PN, EN, RN) is iterated. Within this framework, the generator is capable of receiving dependable gradients that consistently enhance the embedding's effectiveness.

## 4. Experiments Result and Analysis

In this segment of our study, we have undertaken [29] a series of comprehensive ablation experiments aimed at methodically assessing the impact and efficacy of the various design choices incorporated into our model. These experiments are critical for understanding the contribution of individual components and features to the overall performance of our system. By methodically disabling specific elements or altering configurations, we can isolate and identify the value and functionality of each discrete design decision.

### 4.1. Experiment Environment

We delve into the specifics of the datasets used for testing and validation, as well as the details regarding the experimental framework structured for the evaluation of our model. We meticulously select comprehensive datasets, ensuring a broad representation of various image types and complexity to robustly train and test our system. We then elaborate on the configuration of our experimental setup, which includes the hardware specifications, software environments, and the parameters set for conducting the experiments. This will provide clarity on how the trials were performed and under what conditions, setting the stage for replicable and transparent results. In the selection of data sets, we take

ImageNet [31] is extensively utilized as the primary data source for challenges in image recognition and classification. Meanwhile, the NWPU-RESISC45 [32] database, comprising 31,500 aerial images spanning 45 different scene categories with abundant spatial diversity and variation, serves as the repository for the RSI collection.

During our experimental setup, we utilized a workstation equipped with an NVIDIA GeForce RTX 3080 Ti GPU for hardware. The software environment was configured with the Python 3.8 programming language and the Pytorch 1.31.0 framework, operating on the Windows 10 platform. For training purposes, we employed a corpus of 30,000 images from the ImageNet dataset as the cover images, alongside 1,500 images designated for testing. In a similar fashion, for the confidential images, we trained on 30,000 images and allocated 1,500 images for testing, all sourced from the NWPU-RESISC45 dataset.

### 4.2. Visual Quality Test

The main purpose of this scheme is to hide RSI through natural images to achieve efficient and secure distribution. One of the most basic requirements is that the hidden image is not visually detectable by the attack, because if the attack finds the image suspicious, it will steal stego or block the distribution channel. Therefore, the first experiment we did after the program was trained was to test the visual effects of stego. In addition, we also tested the visual quality of the RSI extracted from stego, aiming to test the hidden performance of the scheme from the perspective of visual quality. Figure 1 shows the 6 randomly selected image pairs in our experimental results, which are Cover, RSI, Stego from left to right, the extracted RSI and the residual between Cover and Stego. Intuitively, the stego generated by the scheme and the extracted RSI have good visual effects. Subjective evaluation is not very accurate, and we have made some objective evaluation, mainly using PSNR [33] and SSIM [34] two general visual quality evaluation criteria. Table 1 shows the results of objective evaluation. It is generally believed that PSNR reaches 37db and SSIM reaches 0.85, which means that the evaluation object has better visual quality, that is, no visual anomaly can be detected compared with the reference object.

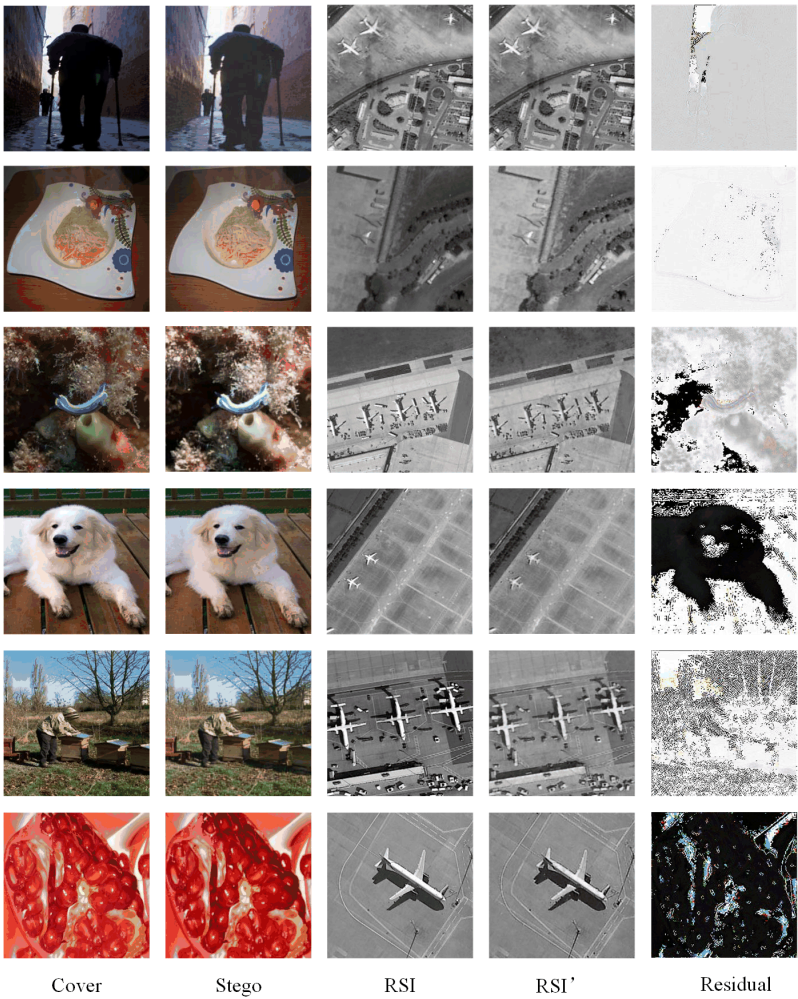


Figure 7. Visual performance display.

Table 2. Objective evaluation of embedding and revealing effects.

The test pairs	Embedding	Revealing
	PSNR/SSIM	PSNR/SSIM
Row#1	46.8db/0.97	38.7db/0.86
Row#2	47.1db/0.98	39.3db/0.84
Row#3	46.9db/0.96	39.2db/0.86
Row#4	46.8db/0.97	38.6db/0.88
Row#5	47.2db/0.98	38.8db/0.86
Row#6	46.9db/0.96	39.1db/0.85

\* PSNR and SSIM can be calculated by equation 18 and 19 respectively.

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX^2}{MSE} \right) \tag{18}$$

In equation (18), *MAX* is the maximum possible pixel value of the two images involved in the calculation and *MSE* represents the Mean Squared Error.

$$SSIM(x,y) = \frac{(2u_xu_y + C_1)}{(u_x^2 + u_y^2 + C_1)} \cdot \frac{(2\delta_{xy} + C_2)}{(\delta_x^2 + \delta_y^2 + C_2)} \tag{19}$$



In equation (19),  $u_x$  and  $u_y$  represent the average values of images,  $\delta_{xy}$  represents covariance between images,  $\delta_x$  and  $\delta_y$  represent the variances of images,  $C_1$  and  $C_2$  are two constants which are used to prevent unstable results.

4.3. Semantic Retention Capability Test

In image classification, three standard evaluation metrics are commonly utilized: overall accuracy, average accuracy, and the confusion matrix. Overall classification accuracy (OCA) is measured by the proportion of correctly classified samples across all classes relative to the total sample count. Average classification accuracy (ACA) calculates the mean classification accuracy for each class, independent of the class sample size. The confusion matrix, an insightful layout, serves to dissect the classification performance, detailing each correct or mistaken prediction by class through an accumulative tabulation of tested samples.

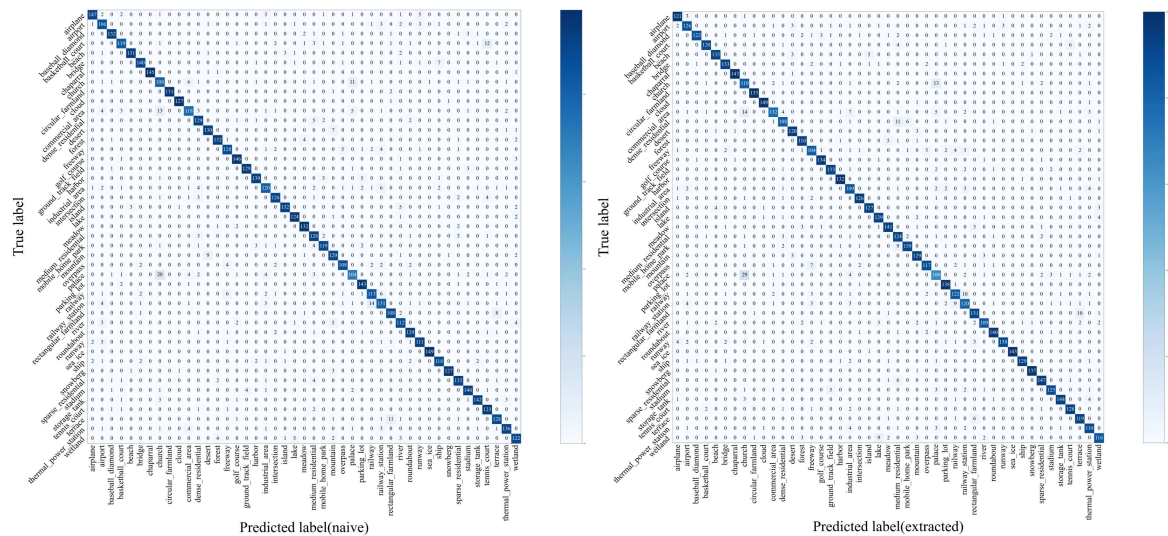


Figure 8. Embedding network architecture.

It is important to note that in the case of the NWPU-RESISC45 dataset, each class contains an identical number of images. Consequently, the overall accuracy coincides with the average accuracy. As a result, in our study, we only employ overall accuracy and the confusion matrix to gauge the effectiveness of various classification methodologies.

Furthermore, to ensure the dependability of the overall accuracy and confusion matrix metrics, we conducted ten iterations of the experimental process for each training-testing split. The outcomes are then presented as a mean and standard deviation, providing a robust and trustworthy statistical analysis of the classification results.

Table 3. Overall Accuracy of three kinds of method based on CNN and their fine-tuned variants under the training ratios of 70% and 80%.

Method based on CNN	70% training ratio	80% training ratio
	Native/extracted	Native/extracted
AlexNet	91.5±0.18/91.3±0.17	92.7±0.12/92.6±0.11
VGGNet16	90.5±0.19/90.6±0.15	92.6±0.20/92.6±0.19
GoolgeLeNet	91.8±0.13/92.1±0.12	92.3±0.13/92.2±0.15
Fine-tuned AlexNet	97.5±0.18/97.2±0.16	98.7±0.10/98.5±0.11
Fine-tuned VGGNet16	97.6±0.18/96.9±0.19	98.9±0.09/98.3±0.08
Fine-tuned GoolgeLeNet	97.3±0.18/97.5±0.17	98.7±0.12/98.4±0.13



#### 4.4. Security Test

Security index is an important measure of the practicality of the scheme in this paper, which mainly refers to the ability of the scheme to resist detection by relevant algorithms, that is, to detect whether stego contains RSI. Referring to the general practice of security testing in steganography, we choose two steganographic analysis algorithms, one is based on statistical analysis and the other is based on deep learning. The former representative tool is stegexpose, and the latter representative algorithm is XuNet. In order to obtain the experimental results accurately, 10,000 natural images from ImageNet were selected in the test, and the corresponding stego was obtained by inputting these 10,000 images into the trained embedded model. First of all, the stegexpose resistance test was carried out, and 2000 images were randomly selected and input into the stegexpose program, and the result was obtained. From the results, the stegexpose detection result is 0.51 similar to the result of random guessing. It can be seen that the proposed algorithm does not affect the least significant bit, DCT coefficient and noise distribution of cover image, and can resist the analysis attack of Stegexpose with drug efficacy. In other words, The scheme in this paper is safe when only the stegexpose test is used. In order to further verify the security of our scheme, we then carried out the experiment of resistance to YeNet, and randomly selected 8000 pairs of the above 10000 pairs of images for training, and then used the remaining 2000 pairs as a test. From the experimental results, the accuracy of YeNet can reach 0.90%. However, this high level does not mean that our algorithm is unsafe, because this supervised detection algorithm needs to obtain the stego generated by our algorithm and the corresponding cover for model training, which is almost impossible in many cases.

#### 4.5. Quantitative Comparison

In order to verify the necessity of this study, we do some comparative experiments to verify that the proposed algorithm is more suitable for our RSI security distribution task from different perspectives. Specifically, we selected six deep information hiding algorithms [8,11,12,20,24,35] that are close to the idea and method of the algorithm in this paper. They only completed the process of hiding pictures in pictures, without special design for RSI. Therefore, in order to make the comparison fair, we only compared some performance indicators that were tested jointly in these papers, including stego visual quality, extracted RSI visual quality, and detection resistance.

In the experiment, we replicated schemes mentioned above and tested them with the same cover and RSI. The experimental results are shown in Table 4. According to the experimental results, the proposed algorithm achieves the best visual effect and anti-YeNet analysis ability, which is mainly due to the PN and PAM module designed in this paper according to the characteristics of RSI and the addition of DN network, which is further verified in the subsequent ablation experiment.

**Table 4.** Performance comparison with existing algorithms.

Schemes	Embedding	Revealing	Accuracy of Detection
	PSNR/SSIM	PSNR/SSIM	YeNet
Literature [8]	34.78db/0.92	31.5db/0.90	0.55/0.99
Literature [21]	34.6db/0.96	36.1db/0.94	0.53/0.98
Literature [11]	44.1db/0.97	39.8db/0.98	0.58/0.98
Literature [20]	41.3db/0.95	33.1db/0.97	0.52/0.98
Literature [24]	44.6db/0.97	38.6db/0.97	0.53/0.98
Literature [12]	42.3db/0.99	38.8db/0.96	0.52/0.96
The proposed method	43.8db/0.99	38.9db/0.99	0.51/0.90

\* PSNR and SSIM can be calculated by equation 18 and 19 respectively.

#### 4.6. Ablation Experiments

In our study, we incorporated the Position Attention Module (PAM) and the Perceptual Network (PN) into our Inception Network-based architecture to dynamically refine the distribution of Remote

Sensing Image (RSI) data. The PN within our framework serves to enhance the embedding ability of the Encoder Network (EN) by progressively approximating the visual characteristics and distribution of the cover images, and to improve the restoration capability of the Revealing Network (RN) by ensuring the recovered images closely match the visual and semantic qualities of the RSI.

To assess the impact of the PAM and PN, we conducted an ablation study with variations to the original DIH4RSID configuration, producing three offshoots: (1) DIH4RSID-PAM-PN lacking both PAM and PN; (2) DIH4RSID-PAM without PAM, but including PN; (3) DIH4RSID-PN absent PN, but incorporating PAM. We trained the main DIH4RSID network and its variants on the same dataset and evaluated their efficacy based on image quality, extraction accuracy, visual appeal, security, and classification precision. The quantitative outcomes are presented in Table 5.

The comparison between DIH4RSID and DIH4RSID-PAM-PN clearly demonstrates the significant enhancements PAM and PN provide across all benchmarks. These components optimize feature utilization, highlighting relevant details while suppressing extraneous ones, thereby generating steganographic images with less perceptible noise. Additionally, the inclusion of PN in DIH4RSID aids in more accurate recovery of RSIs.

A marginal decline in performance with DIH4RSID-PN suggests that PAM contributes to stable embedding by promoting the imperceptibility of steganographic content. Furthermore, the effectiveness of RSI extraction and visual quality also relies on PN, as it facilitates the concurrent training of both the EN and RN, underscoring the complementary roles of PAM and PN in our network design.

Table 5. Ablation study results.

Variants	Embedding	Revealing	AD	ACA
	PSNR/SSIM	PSNR/SSIM	Stegexpose/YeNet	
DIH4RSID-PAM-PN	36.8db/0.80	29.6db/0.75	0.53/0.96	0.89
DIH4RSID-PAM	40.9db/0.83	30.1db/0.78	0.55/0.92	0.94
DIH4RSID-PN	42.1db/0.92	32.3db/0.80	0.52/0.90	0.93
DIH4RSID	47.1db/0.96	38.3db/0.84	0.53/0.85	0.98

\* AD denotes as the accuracy of detection of Stegexpose or YeNet and ACA denotes as the average accuracy classification.

5. Discussion

In order to provide an alternative scheme for the safe distribution of remote sensing images, this paper proposes a new end-to-end network structure based on the idea of deep information hiding. In fact, the remote sensing images we want to hide are essentially no different from ordinary natural images, so an intuitive idea is to transplant the existing image-to-hide image algorithm, and through comparative experiments and ablation, it can be found that in fact, it can be barely used in scenes with low performance requirements. However, further experimental results show that the proposed algorithm can achieve better visual quality and higher security. Through analysis, it can be seen that the scheme in this paper achieves better performance than the existing algorithms because of the delicate design of network structure.

Firstly, this paper designs a PN according to the characteristics of RSI. Theoretically, in order to capture more details of RSI and targets of different scales, the features extracted by PN module are used as input for subsequent embedding in the network. Secondly, a newly designed Attention mechanism module, PAM, is adopted in the encoder network, which is similar to Coordinate Attention Machnism, which can realize feature enhancement and context information integration. In addition to adding PAM to the encoder, in order to achieve good information embedding and visual quality maintenance, we also add Local skip connection and Global skip connection. Finally, in order to achieve better security, we add the discriminator module to the pipeline to achieve higher security and embedding effect.

Although this study provides a novel approach for the secure distribution of remote sensing images, the RSI images used are compressed RGB images, thus the method presented in the paper cannot be directly applied to hyperspectral remote sensing images. Therefore, to extend the applicability of this algorithm, future work could attempt to implement the embedding and extraction for hyperspectral images to achieve their secure distribution.

**Author Contributions:** Conceptualization, P.L., J.L. and D.J.M.; methodology, P.L., J.L.; validation, P.L., J.L. and D.J.M.; formal analysis, P.L., J.L.; investigation, P.L.; writing—original draft preparation, P.L., J.L.; writing—review and editing, P.L., J.L.; experiments, P.L.; supervision, P.L., J.L. and D.J.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** his work was supported in part by the National Key R&D Program of China under grant 2021YFB3100901, NSF of China under Grant 62074131, 62272389, 62372069 and Shaanxi Provincial Key R&D Program 2023-ZDLGY-32.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Zhang, D.; Ren, L.; Shafiq, M.; Gu, Z. A Lightweight Privacy-Preserving System for the Security of Remote Sensing Images on IoT. *Remote Sensing* **2022**, *14*. <https://doi.org/10.3390/rs14246371>.
2. Zhang, X.; Zhang, G.; Huang, X.; Poslad, S. Granular Content Distribution for IoT Remote Sensing Data Supporting Privacy Preservation. *Remote Sensing* **2022**, *14*. <https://doi.org/10.3390/rs14215574>.
3. Alsubaei, F.S.; Alneil, A.A.; Mohamed, A.; Mustafa Hilal, A. Block-Scrambling-Based Encryption with Deep-Learning-Driven Remote Sensing Image Classification. *Remote Sensing* **2023**, *15*. <https://doi.org/10.3390/rs15041022>.
4. Naman, S.; Bhattacharyya, S.; Saha, T. Remote sensing and advanced encryption standard using 256-Bit key. In Proceedings of the Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph 2018. Springer, 2020, pp. 181–190.
5. He, R.; Sun, Q.; Thangasamy, P.; Chen, X.; Zhang, Y.; Wang, H.; Luo, H.; Zhou, X.D.; Zhou, M. Accelerate oxygen evolution reaction by adding chemical mediator and utilizing solar energy. *International Journal of Hydrogen Energy* **2023**, *48*, 8898–8908.
6. Akhaee, M.A.; Marvasti, F. A Survey on Digital Data Hiding Schemes: Principals, Algorithms, and Applications. *ISeCure* **2013**, *5*.
7. Singh, A.K. Data hiding: current trends, innovation and potential challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **2020**, *16*, 1–16.
8. Rehman, A.; Rahim, R.; Nadeem, M.; Hussain, S. End-to-end trained CNN encode-decoder networks for image steganography. In Proceedings of the Proc. Comput. Vis.-ECCV 2018 Workshops, 2018, pp. 723–729.
9. Zhang, K.A.; Cuesta-Infante, A.; Xu, L.; Veeramachaneni, K. SteganoGAN: High Capacity Image Steganography with GANs. *arxiv* **2019**.
10. Yu, C. Attention Based Data Hiding with Generative Adversarial Networks. In Proceedings of the Proceedings of the AAAI conference on artificial intelligence, 2020, pp. 1120–1128.
11. Chen, F.; Xing, Q.; Liu, F. Technology of hiding and protecting the secret image based on two-channel deep hiding network. *IEEE Access* **2020**, *8*, 21966–21979.
12. Chen, F.; Xing, Q.; Fan, C. Multilevel Strong Auxiliary Network for Enhancing Feature Representation to Protect Secret Images. *IEEE Transactions on Industrial Informatics* **2022**, *18*, 4577–4586.
13. Mielikainen, J. LSB matching revisited. *IEEE signal processing letters* **2006**, *13*, 285–287.
14. Filler, T.; Judas, J.; Fridrich, J. Minimizing Additive Distortion in Steganography Using Syndrome-Trellis Codes. *IEEE Transactions on Information Forensics and Security* **2011**, *6*, 920–935. <https://doi.org/10.1109/TIFS.2011.2134094>.
15. Pevný, T.; Filler, T.; Bas, P. Using high-dimensional image models to perform highly undetectable steganography. In Proceedings of the Information Hiding: 12th International Conference, IH 2010, Calgary, AB, Canada, June 28–30, 2010, Revised Selected Papers 12. Springer, 2010, pp. 161–177.

16. Holub, V.; Fridrich, J. Designing steganographic distortion using directional filters. In Proceedings of the 2012 IEEE International Workshop on Information Forensics and Security (WIFS), 2012, pp. 234–239. <https://doi.org/10.1109/WIFS.2012.6412655>.
17. Holub, V.; Fridrich, J. Digital image steganography using universal distortion. In Proceedings of the Proceedings of the first ACM workshop on Information hiding and multimedia security. ACM, 2013, pp. 59–68.
18. Li, B.; Wang, M.; Huang, J.; Li, X. A new cost function for spatial image steganography. In Proceedings of the 2014 IEEE International conference on image processing (ICIP). IEEE, 2014, pp. 4206–4210.
19. Zhu, J.; Kaplan, R.; Johnson, J.; Fei-Fei, L. Hidden: Hiding data with deep networks. In Proceedings of the Proceedings of the European conference on computer vision (ECCV), 2018, pp. 657–672.
20. Baluja, S. Hiding Images within Images. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **2020**, 42, 1685–1697.
21. Duan, X.; Jia, K.; Li, B.; Guo, D.; Zhang, E.; Qin, C. Reversible image steganography scheme based on a U-Net structure. *IEEE Access* **2019**, 7, 9314–9323.
22. Huang, J.; Luo, T.; Li, L.; Yang, G.; Xu, H.; Chang, C.C. ARWGAN: Attention-guided Robust Image Watermarking Model Based on GAN. *IEEE Transactions on Instrumentation and Measurement* **2023**.
23. Tan, J.; Liao, X.; Liu, J.; Cao, Y.; Jiang, H. Channel attention image steganography with generative adversarial networks. *IEEE transactions on network science and engineering* **2021**, 9, 888–903.
24. Chen, F.; Xing, Q.; Sun, B.; Yan, X.; Cheng, J. An Enhanced Steganography Network for Concealing and Protecting Secret Image Data. *Entropy* **2022**, 24, 1203.
25. Shi, J.; Liu, W.; Shan, H.; Li, E.; Li, X.; Zhang, L. Remote Sensing Scene Classification Based on Multibranch Fusion Attention Network. *IEEE Geoscience and Remote Sensing Letters* **2023**, 20, 1–5. <https://doi.org/10.1109/LGRS.2023.3262407>.
26. Xu, G.; Wu, H.Z.; Shi, Y.Q. Structural Design of Convolutional Neural Networks for Steganalysis. *IEEE Signal Processing Letters* **2016**, 23, 708–712. <https://doi.org/10.1109/LSP.2016.2548421>.
27. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2018**, 40, 834–848. <https://doi.org/10.1109/TPAMI.2017.2699184>.
28. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in neural information processing systems, 2014, Vol. 27.
29. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein generative adversarial networks. In Proceedings of the Proceedings of the 34th International Conference on Machine Learning-Volume 70. JMLR. org, 2017, pp. 214–223.
30. Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A. Improved Training of Wasserstein GANs, 2017.
31. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. Imagenet large scale visual recognition challenge. *International journal of computer vision* **2015**, 115, 211–252.
32. Cheng, G.; Han, J.; Lu, X. Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE* **2017**, 105, 1865–1883.
33. Almohammad, A.; Ghinea, G. Stego image quality and the reliability of PSNR. In Proceedings of the International Conference on Image Processing, 2010.
34. Wang, Z.; Bovik, A.; Sheikh, H.; Simoncelli, E. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **2004**, 13, 600–612. <https://doi.org/10.1109/TIP.2003.819861>.
35. Zhang, R.; Dong, S.; Liu, J. Invisible steganography via generative adversarial networks. *Multimedia tools and applications* **2019**, 78, 8559–8575.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.