

Article

Progressive Transmission Line Image Transmission and Recovery Algorithm Based on Hybrid Attention and Feature Fusion for Signal-Free Regions of Transmission Lines

Xiu Ji ^{1,*}, Xiao Yang ², Zheyu Yue ², Hongliu Yang ³ and Haiyang Guo ²

¹ Institute of Future Industrial Technology Innovation, Changchun Institute of Technology, Changchun 130012, China

² School of Electrical and Electronic Engineering, Changchun University of Technology, Changchun 130022, China; 2202304113@stu.ccut.edu.cn (X.Y.); 20231041225@stu.ccut.edu.cn (Z.Y.); 2202404083@stu.ccut.edu.cn (H.G.)

³ State Grid Jilin Electric Power Co., Ltd., Tonghua Power Supply Company, Tonghua 134100, China; 2202304091@stu.ccut.edu.cn

* Correspondence: jixiu523@ccit.edu.cn

Abstract: In this paper, a progressive image transmission and recovery algorithm based on hybrid attention mechanism and feature fusion is proposed, aiming to solve the challenge of monitoring the signal-less region of transmission lines. The method combines wavelet transform, Swin Transformer, and hybrid attention module with the Pixel Shuffle upsampling mechanism to achieve a balance between quality and efficiency of image transmission in a low bandwidth environment. Initial preview is achieved by prioritizing the transmission of low-frequency subbands through wavelet transform, followed by dynamic optimization of the weight allocation of key features using a hybrid attention and local window multi-scale self-attention mechanism, and further enhancement of the resolution of the decoded image through Pixel Shuffle upsampling. Experimental results show that the algorithm significantly outperforms existing methods in terms of image quality (PSNR, SSIM), transmission efficiency, and bandwidth utilization, proving its superior adaptability and effectiveness in surveillance scenarios in signal-free regions.

Keywords: progressive image transmission; wavelet transform; hybrid attention mechanism; Swin Transformer; Pixel Shuffle upsampling module



Citation: Ji, X.; Yang, X.; Yue, Z.; Yang, H.; Guo, H. Progressive Transmission Line Image Transmission and Recovery Algorithm Based on Hybrid Attention and Feature Fusion for Signal-Free Regions of Transmission Lines. *Electronics* **2024**, *13*, 4605. <https://doi.org/10.3390/electronics13234605>

Academic Editor: Stanley C. Ahalt

Received: 31 October 2024

Revised: 17 November 2024

Accepted: 20 November 2024

Published: 22 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the continuous expansion of power systems and the growing coverage of transmission lines, ensuring the stable operation and reliable monitoring of transmission infrastructure has become a critical issue in the power industry. However, many transmission lines are located in remote areas, such as mountains and wilderness, where communication signals are unavailable, making it particularly challenging to obtain real-time monitoring images. The presence of these signal-free areas not only hampers fault detection and emergency response, but also threatens the stable operation of the power grid. Consequently, achieving efficient image transmission and progressive recovery in areas with limited communication has become a key challenge in the operation and maintenance of power grids [1].

Traditional surveillance systems typically rely on fixed or wireless communication networks for image transmission. However, in low-bandwidth or completely signal-free environments, these conventional methods often fail to meet surveillance requirements due to slow transmission speeds and high latency. While previous image compression and transmission techniques can partially alleviate bandwidth limitations, they struggle to preserve image quality and retain essential details—particularly in complex natural environments. The loss of crucial image information can delay fault detection, leading to

increased maintenance costs and operational risks. Therefore, there is a pressing need for an algorithm that can not only transmit images efficiently, but also progressively enhance image quality throughout the transmission process.

To address the above challenges, this paper proposes a progressive image transmission algorithm based on a hybrid attention mechanism. The algorithm effectively balances image quality and transmission efficiency under low-bandwidth conditions by integrating wavelet transform, the Swin Transformer, a hybrid attention mechanism, and the Pixel Shuffle upsampling module. Specifically, the wavelet transform decomposes the image into low- and high-frequency subbands, prioritizing the transmission of low-frequency subbands to provide a fast preview, thereby reducing unnecessary data transfer. The hybrid attention mechanism combines channel attention and pixel attention modules to dynamically adjust the weights of key image features, enhancing detail recovery. The Swin Transformer captures both global and local features through a local window multi-scale self-attention mechanism, improving the model's overall feature comprehension. Finally, the Pixel Shuffle upsampling module enhances resolution by rearranging channel data, ensuring the decoded image is of high quality with rich details.

Compared to existing algorithms, the proposed algorithm significantly improves image quality (PSNR, SSIM), transmission efficiency, and bandwidth utilization, making it better suited for monitoring needs in extreme environments. This paper verifies the effectiveness of the algorithm through experiments conducted in various scenarios. The results demonstrate that the method delivers excellent performance and stability in monitoring transmission lines in signal-free areas. Additionally, this paper outlines the following key contributions.

- (1) A progressive image transmission and recovery algorithm based on the hybrid attention mechanism is proposed, which achieves a balance between image quality and transmission efficiency in low bandwidth environments and is particularly suitable for transmission line monitoring in signal-free regions.
- (2) The wavelet transform is innovatively combined with Swin Transformer, hybrid attention module, and pixel rearrangement upsampling mechanism to optimize the feature extraction and image recovery process, which significantly improves the clarity and detail retention of the recovered images.
- (3) Experimental results show that the algorithm performs superiorly in terms of image quality (PSNR, SSIM), transmission efficiency, and bandwidth utilization compared with mainstream methods, demonstrating its excellent adaptability and stability in extreme environment monitoring scenarios.

2. Related Work

2.1. Approaches Based on Attention Mechanisms

In recent years, numerous studies have explored attention mechanisms to enhance the efficiency and quality of image processing. Wang et al. [2] introduced an efficient channel attention (ECA) module that employs one-dimensional convolution to facilitate cross-channel interactions and adaptively adjusts the convolution kernel size to improve performance. This method achieves notable performance gains with minimal parameter overhead but fails to address both spatial and channel information simultaneously. In response, our hybrid attention mechanism integrates channel- and pixel-level attention to comprehensively capture image features and optimize image reconstruction in low-bandwidth environments. Additionally, efficient image transmission is vital in medical imaging for telemedicine applications [3], particularly under constrained bandwidth conditions. While existing research employs compression-aware and deep learning-based compression strategies to maintain diagnostic accuracy, our work advances bandwidth efficiency and ensures higher transmission performance through progressive recovery techniques.

The Residual Channel Attention Network (RCAN) developed by Zhang et al. [4] utilizes a residual-in-residual (RIR) structure that bypasses low-frequency information via multiple skip connections, concentrating on high-frequency feature learning. While RCAN

excels at preserving fine details, its high computational complexity limits adaptability across varying bandwidth constraints. Our approach mitigates this complexity by incorporating Swin Transformer and Pixel Shuffle upsampling modules, enhancing both global and local feature comprehension. This makes it particularly effective for transmission line monitoring in signal-blind areas. Similar to the bandwidth limitations encountered in satellite remote sensing [5], which employs hierarchical image coding and lossless compression to prioritize essential information transmission, our method demonstrates superior adaptability for handling complex textures and maintaining detailed information.

The SCA-CNN developed by Chen et al. [6] integrates spatial and channel attention mechanisms to improve feature selection for image description tasks. However, this approach is not designed for progressive image recovery in bandwidth-constrained environments. By contrast, our method leverages the wavelet transform in conjunction with the Swin Transformer module to enable efficient image transmission and reconstruction in signal-blind areas using progressive transmission and recovery techniques. This approach aligns with the requirements of disaster management and emergency response [7], where low-bandwidth radio networks and progressive image reconstruction are employed to address urgent situations. Our algorithm further enhances image quality and transmission efficiency in these critical scenarios.

2.2. Image Restoration Technology

In the field of image restoration, Zhao et al. [8] developed a lightweight convolutional network featuring a pixel attention (PA) mechanism to enhance image reconstruction quality through self-calibrating convolution and efficient upsampling. Although this method achieves strong performance with minimal parameter overhead, it struggles with preserving complex textures. Our algorithm addresses this challenge by combining a hybrid attention mechanism with a window-based multi-scale self-attention strategy, enhancing image clarity across various scenes. Similarly, in applications such as autonomous driving and robotics, where real-time image data transmission is crucial under limited bandwidth conditions [9], edge computing and feature-prioritized transmission methods are commonly employed. Our approach further optimizes image recovery and transmission by leveraging multi-scale information extraction and upsampling techniques.

Xiao et al. [10] proposed a multi-feature selection method for small target detection in complex backgrounds, enhancing detection performance through a bidirectional multi-scale feature fusion network. However, this method primarily addresses target detection rather than progressive image recovery. In contrast, our research focuses on enhancing image transmission efficiency and recovery quality in signal-blind surveillance scenarios. Related studies have utilized wavelet transform and compression algorithms to optimize image transmission in underwater and space detection. Our approach, however, further refines detail recovery in complex environments by employing pixel shuffling upsampling and advanced attention mechanisms.

The SwinIR model developed by Liang et al. [11] demonstrates outstanding performance in denoising and artifact removal tasks. However, its high computational resource requirements limit its applicability in bandwidth-constrained scenarios. Similarly, the Hierarchical Swin Transformer (HST) proposed by Li et al. [12] effectively captures hierarchical features but also faces significant resource consumption challenges. Our approach offers a more efficient solution for low-bandwidth environments by prioritizing the transmission of essential subbands and employing the Pixel Shuffle module for optimized upsampling.

The single-image super-resolution (SISR) method proposed by Nascimento et al. [13] combines the Pixel Shuffle and attention mechanisms, achieving strong performance in static scenes but lacking the capability to handle dynamic transmission conditions. Our approach extends Pixel Shuffle within an asymptotic recovery framework, ensuring high adaptability for real-world surveillance tasks.

3. The Structure of the Model Architecture

3.1. Wavelet Transform

Wavelet transform is a multiresolution analysis tool used to decompose an image signal into low- and high-frequency subbands, allowing the extraction of various levels of image information [14]. The low-frequency subbands primarily capture the global structure and most of the image's energy, serving as a rough representation of the image, while the high-frequency subbands retain details and edge information. By processing these frequency components, wavelet transform enables effective progressive recovery during image transmission. In each wavelet transform, the image is decomposed into one low-frequency and several high-frequency subbands. After N levels of wavelet transformation, the image is divided into a low-frequency subbands (LL_N) and multiple high-frequency subbands (HL_i, LH_i, HH_i), where $i = 1, 2, 3, \dots, N$. The low-frequency subbands contain most of the image's energy, providing the main content and structural information, while the high-frequency subbands capture details such as edges and textures.

According to the energy analysis of different frequency subbands, the low-frequency subbands contain most of the image's energy. Table 1 illustrates this using the classical Lena image as an example [15], where the low-frequency subbands account for over 90% of the total energy, while the energy proportion in the high-frequency subbands decreases progressively across levels. The high-frequency subbands are further divided into horizontal (HL_i), vertical (LH_i), and diagonal (HH_i) components, each capturing specific structural details of the image. Typically, the high-frequency subbands in the vertical and horizontal directions represent prominent features, such as edges and contours, making their information particularly significant.

Table 1. Statistics of the energy ratios for each wavelet subband of the Lena image.

Resolution Levels	Wavelet Sub-Band ub-Band	Energy Ratio (%)	The Sub-Band Energy Ratio is the Sum of the Total
low frequency	LL_4	91.6367	91.6367
Layer 4 HF ($LH_4+HL_4+HH_4$)	LH_4	0.7001	3.896
	HL_4	2.5976	
	HH_4	0.3913	
Layer 3 HF ($LH_3+HL_3+HH_3$)	LH_3	0.5021	2.3257
	HL_3	1.5530	
	HH_3	0.2706	
Layer 2 HF ($LH_2+HL_2+HH_2$)	LH_2	0.3473	1.4489
	HL_2	0.9484	
	HH_2	0.1532	
Layer 1 HF ($LH_1+HL_1+HH_1$)	LH_1	0.2302	0.8998
	HL_1	0.5696	
	HH_1	0.1001	

During image reconstruction, the low-frequency subbands are first used to generate a low-resolution version of the image through wavelet inversion. As the subsequent high-frequency subbands are progressively transmitted and decoded, the image details are incrementally restored, resulting in a continuous improvement in image quality [16]. This gradual recovery of image details not only reduces the bandwidth requirements for the initial transmission, but also allows users to assess whether the image meets their needs at the low-resolution stage, thereby minimizing unnecessary data transmission.

The energy of each subband is calculated by summing the squares of all its coefficients, as expressed in the following formula:

$$E_{sub-band} = \sum_{i,j} c_{i,j}^2$$

The statistical results in Table 1 show that the low-frequency subbands contain most of the image's energy, while the high-frequency subbands account for only a small portion. Additionally, as the resolution decreases, the proportion of energy in the high-frequency subbands increases. Even at the same resolution, the energy distribution varies among the high-frequency subbands: the vertical direction (HLi) contains the highest energy, followed by the horizontal direction (LHi), with the diagonal direction (HHi) containing the least. During the reconstruction stage, the system first generates a low-resolution preview image using the low-frequency subbands. As the high-frequency subbands are gradually decoded, the image details are progressively restored, leading to continuous improvement in image quality. This approach ensures that essential details are incrementally recovered, even under limited bandwidth conditions.

3.2. Hybrid Attention Module

To enhance performance during feature extraction and ensure the network effectively filters out unimportant image regions while retaining key details, we designed a hybrid attention module, as illustrated in Figure 1 [17]. The blue section on the left represents the pixel attention (PA) mechanism, which assigns weights at the pixel level by capturing the local spatial relationships between pixels [18]. For pixels containing critical details, PA assigns higher weights to enhance texture and detail representation in the reconstructed image, while for less important pixels, it assigns lower weights, effectively reducing noise and preserving essential information. The right section depicts the channel attention (CA) mechanism, which emphasizes important feature channels and suppresses less relevant ones by dynamically adjusting their weights [19]. At the top is the blending module, which multiplies the pixel-level and channel-level correlation matrices and adds the result to the input feature map. This mechanism enables the network to focus on the most critical channel features during image reconstruction, improving its comprehension of the overall image content. By incorporating multiple attention mechanisms, this module enhances the network's ability to capture essential image features, generating clearer and more accurate reconstructed images.

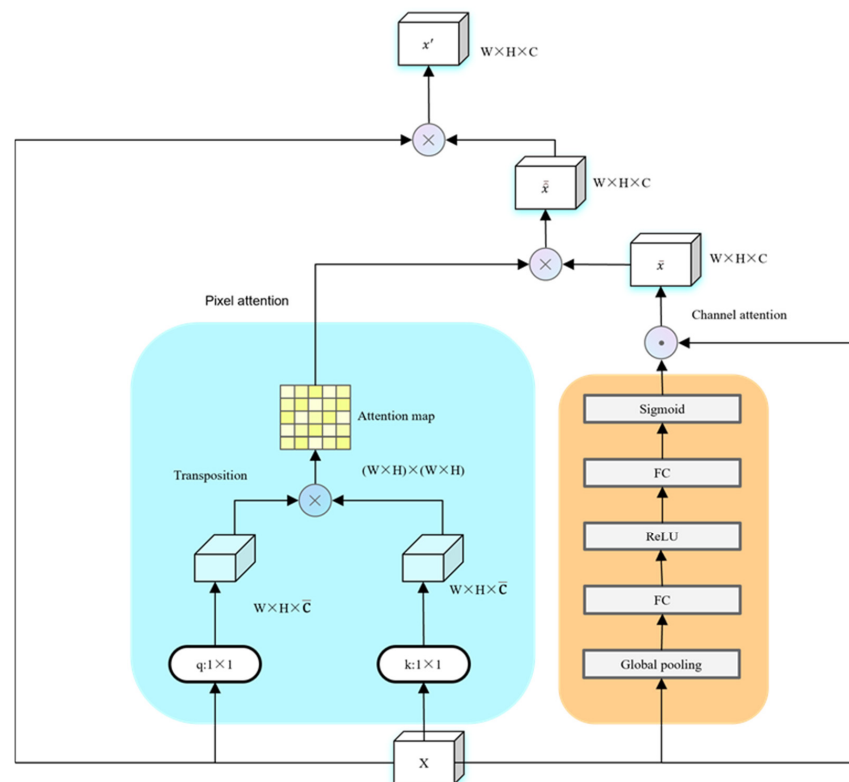


Figure 1. Hybrid attention module.

The hybrid attention module combines channel attention (CA) and pixel attention (PA) to focus on key information within an image in a multi-layered manner, effectively enhancing the representation of image details. It employs a residual connection structure to seamlessly integrate the CA and PA mechanisms into a unified and efficient framework. The specific calculation process is outlined as follows:

$$F_i = PA(CA(Conv(Conv(F_{i-1}) + F_{i-1}))) + F_{i-1}$$

Let F_{i-1} denote the input to the i th hybrid attention module, $CA(\cdot)$ represent the channel attention operation, and $PA(\cdot)$ the pixel attention operation. The calculation process begins with F_{i-1} undergoing a convolution operation, followed by a residual summation with the original input F_{i-1} . The resulting feature is then processed through another convolution operation, followed sequentially by the channel attention and pixel attention operations. Finally, the output feature F_i is obtained through a residual summation with F_{i-1} .

3.3. Swin Transformer Module

The Swin Transformer incorporates a windowed multi-scale self-attention mechanism to capture global image information, enhancing its ability to extract local features and spatial relationships, and thereby improving the capture of detailed information [20]. This mechanism allows the network to focus on key features within smaller localized areas of polar targets, preserving useful information that facilitates more accurate image recovery. At the same time, it effectively reduces model complexity and improves performance, enabling the network to handle larger-scale images or more complex tasks with the same computational resources.

The feed-forward network, based on a multilayer perceptron (MLP) in the Swin Transformer, has limitations in capturing local context. However, in the field of image recovery, neighboring pixels serve as essential reference points, and convolutional neural networks (CNNs) excel at extracting local features, effectively capturing local information and texture details [21]. To address these limitations, the Swin Transformer is combined with convolutional operations to form a feed-forward Swin Transformer module, as illustrated in Figure 2. This module leverages the windowed self-attention mechanism and hierarchical visual Transformer structure to extract key features from target images in unsigned regions and capture contextual information at multiple scales. By achieving a seamless fusion of global and local information, it enhances the network's ability to comprehend and represent image features. As a result, the network captures structural and textural information more effectively, aiding in the recovery of missing or damaged parts, improving image recovery accuracy, and accelerating processing. Additionally, it maintains strong generalization ability, enabling consistent performance across diverse tasks.

Figure 2a illustrates the Swin Transformer structure, which is computed as follows:

$$\begin{aligned}\hat{Z}_n &= W - MSA(LN(x_{n-1})) + x_{n-1} \\ x_n &= Feed\left(LN\left(\hat{Z}_n\right)\right) + \hat{Z}_n \\ \hat{Z}_{n+1} &= SW - MSA(LN(x_n)) + x_n \\ x_{n+1} &= Feed\left(LN\left(\hat{Z}_{n+1}\right)\right) + \hat{Z}_{n+1}\end{aligned}$$

Let x_{n-1} denote the input to the Swin Transformer module, LN represent the layer normalization operation, W - $MSA(\cdot)$ the window-based multi-head self-attention operation, \hat{Z}_n and \hat{Z}_{n+1} the intermediate features, $Feed(\cdot)$ the feed-forward neural network, and SW - $MSA(\cdot)$ the multi-head self-attention operation based on a moving window. The output of the Swin Transformer is denoted as x_n .

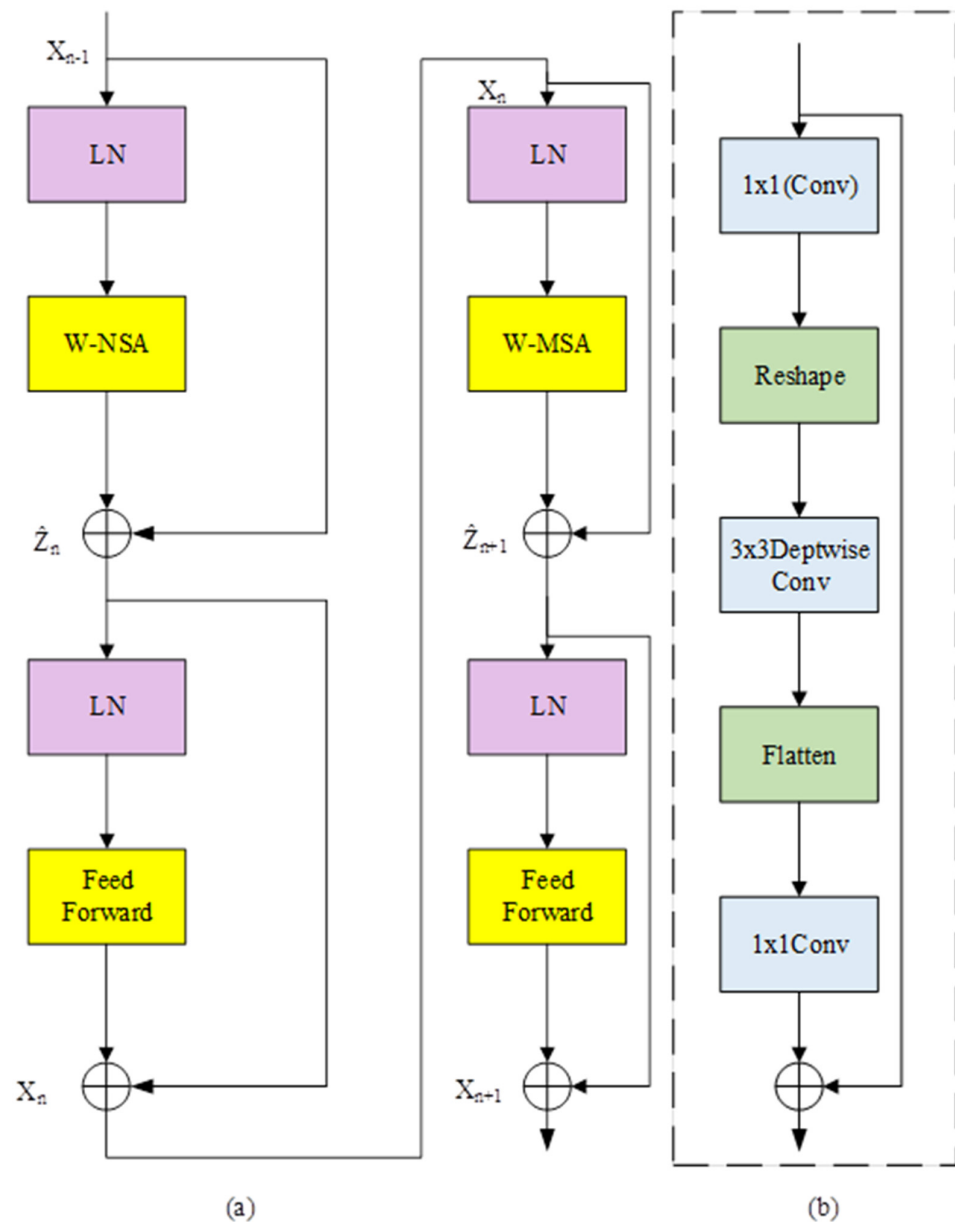


Figure 2. Swin Transformer module (a) Swin Transformer structure (b) Feed Forward structure.

The calculation process begins by normalizing the input feature x_{n-1} using layer normalization. Then, a window-based multi-head self-attention operation produces the intermediate feature \hat{Z}_n by adding the residuals of x_{n-1} . Next, layer normalization, followed by the feed-forward neural network and a residual operation, yields the output of the first block x_n . In the second block, x_n is normalized using layer normalization and processed by the moving window-based multi-head self-attention operation, producing the intermediate feature \hat{Z}_{n+1} through residual addition with x_n . Finally, the second block applies layer normalization, the feed-forward neural network, and a residual operation to produce the final output feature x_{n+1} .

Figure 2b illustrates the structure of the feed-forward neural network, which is computed as follows:

$$x_{n+1} = Conv(Flatten(Dept - wiseConv(Reshape(Conv(x_n))))) + x_n$$

Let x_n denote the input features of the module, $Conv(\cdot)$ the 1×1 convolution operation, $Reshape(\cdot)$ the reshaping of 1D features into 2D features, $Dept-wiseConv(\cdot)$ the

depthwise separable convolution, and Flatten(\cdot) the transformation of 2D features back into 1D features.

The computation begins by altering the feature dimension of x_n using a 1×1 convolution. The resulting feature is then reshaped into a 2D feature map, and a 3×3 depthwise convolution is applied to capture local information. Afterward, the feature is flattened into 1D form, and a 1×1 convolution adjusts the feature dimension to match that of the input. Finally, the output is summed with the residuals of the input feature x_n to produce the final output feature x_{n+1} .

3.4. Pixel Shuffle Upsampling Module

Pixel Shuffle is an upsampling method that rearranges the channel data of a feature map, commonly used in super-resolution and image restoration tasks [22]. By reorganizing multiple channels from a low-resolution feature map into a high-resolution feature map, it significantly enhances detail retention and edge sharpness. In this algorithm, Pixel Shuffle upsampling is employed to optimize the features produced by the feed-forward Swin Transformer module, thereby improving the accuracy of image restoration.

Pixel Shuffle increases the resolution of a feature map by rearranging its depth dimension (number of channels) into the spatial dimensions (width and height). Compared to traditional interpolation methods, Pixel Shuffle better preserves detail information and reduces the occurrence of checkerboard artifacts and other distortions [23].

The specific computation process of Pixel Shuffle upsampling is as follows:

$$y_{n+1} = \text{Conv}(\text{PixelShuffle}(y_n))$$

Let y_n denote the input feature map, Pixel Shuffle denote the Pixel Shuffle operation, and Conv(\cdot) the 1×1 convolution operation. The Pixel Shuffle operation remaps the number of channels in the feature map from $C \times r^2$ to the spatial dimensions, thereby increasing the resolution. The convolution operation further enhances the representational capacity of the feature map after upsampling.

The specific design flow of the Pixel Shuffle upsampling module is illustrated in Figure 3. First, the features produced by the feed-forward Swin Transformer module are used as input. Assume that the dimensions of the input feature maps are $(H, W, C \times r^2)$, where H and W represent the height and width of the input feature maps, respectively. C is the number of channels in the final output, and r is the amplification factor for upsampling. The initial number of channels, $C \times r^2$, contains the additional dimensional information needed to scale the feature map to the target resolution.

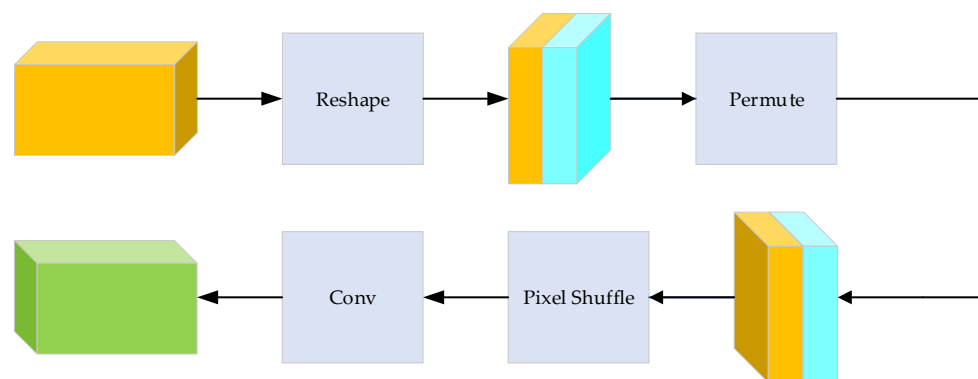


Figure 3. Pixel Shuffle upsampling module.

Second, channel rearrangement and mapping are performed to divide the channels of the input feature map into $r \times r$ subgroups, each containing C channels. This process produces $r \times r$ subfeature maps, each of size (H, W, C) . These subfeature maps correspond to different regions of the target high-resolution feature map. The channel data of each subfeature map are redistributed into the spatial dimensions, resulting in an upsampling

of the feature map's size from (H, W) to $(H \times r, W \times r)$, while the number of channels is reduced to C .

Finally, reorganization and convolution operations are applied to further extract the detailed information of the image. The convolution layer smooths the upsampling results and fuses local information, enhancing the recovery of image details.

By mapping the channel information of the feature map to a higher spatial resolution, Pixel Shuffle effectively preserves the image's details and edge information. This rearrangement prevents the blurring effect often caused by traditional interpolation methods, resulting in clearer texture features.

3.5. Network Infrastructure

The wavelet transform concentrates the main content and energy of the image in the low-frequency subband. Although this subband is crucial for image reconstruction, it contains a relatively small number of coefficients, which decreases inversely with the number of wavelet transform levels, N . After a N -level wavelet transform, the coefficients in the low-frequency subband comprise only $1/4N$ of all the coefficients, and the resolution of the image in this subband is also reduced to $1/4N$ of the original image's resolution. Therefore, in the progressive image transmission process, the lowest-resolution image (an approximation of the original image) can be transmitted and displayed first, allowing users to preview the overall content at low resolution. Based on this preview, users can decide whether higher-resolution and higher-quality image information is needed.

If the image is not required by the user, the transmission of image data can be halted, allowing other image information to be acquired, effectively saving bandwidth. If further improvements to the image quality are needed, the transmission prioritizes subbands containing important information, filtered through the hybrid attention mechanism module. These subbands are processed by the feed-forward Swin Transformer module, which divides them into small, localized regions (windows) and performs self-attention computations within these regions. This approach significantly reduces computational complexity while preserving localized information.

Initially, these processed subbands are represented as low-resolution feature maps. However, the Pixel Shuffle upsampling module rearranges the channel data of these low-resolution feature maps into spatial dimensions (height and width), enhancing their resolution. This ensures that key information in the image reaches a higher quality. Once the key parts achieve satisfactory subjective quality—such that no obvious differences from the original image are perceptible—the process shifts to transmitting subbands containing less important background information. The goal is to use the limited bandwidth efficiently, improving the quality of the background without compromising the overall subjective visual effect, thereby enhancing the image's overall appearance.

After the image undergoes wavelet transformation, the main content and energy are concentrated in the low-frequency subband. Although this subband is essential for image reconstruction, it contains relatively few coefficients and has low resolution, which decreases inversely with the number of wavelet decomposition levels, N . Specifically, after an N -level wavelet transform, the low-frequency subband contains only $1/4N$ of all the coefficients, and the image resolution is reduced to $1/4N$ of the original image.

Based on this, it is a reasonable strategy in progressive image transmission to transmit the image with the lowest resolution first—an approximate representation of the original image after wavelet transformation. This approach allows users to preview the general content of the image and decide whether to continue acquiring higher-resolution versions.

If the user determines that the image does not contain the desired content, the transmission can be halted immediately to save bandwidth. However, if a higher-resolution image is required, the transmission can proceed by prioritizing the important high-frequency information to enhance the image quality.

To optimize the efficiency of subsequent transmissions, a hybrid attention mechanism module can prioritize the transmission of high-frequency subbands containing important

information. The hybrid attention module integrates channel attention and pixel attention mechanisms [24], dynamically adjusting the weights of channels and pixels to focus on enhancing critical details. This process ensures that the subbands most beneficial for improving image quality are prioritized for transmission.

Next, the filtered subbands are processed through the Swin Transformer module. The Swin Transformer divides the image into chunks using a local window self-attention mechanism, which captures local information (e.g., edges, textures) while reducing computational complexity. This approach allows the model to retain critical local details and significantly enhance the efficiency of image recovery.

Although the processed subbands are not particularly sharp, they remain low-resolution feature maps. However, the resolution can be enhanced at this stage using the Pixel Shuffle upsampling module. Pixel Shuffle improves the resolution of the feature map by rearranging its channel information into spatial dimensions (height and width), enabling further enhancement of the key details in the image.

The system begins transmitting the remaining background information only after the critical portions of the image have reached sufficient subjective visual quality—typically when no significant differences from the original image are perceptible. This design ensures that the quality of the most important parts is prioritized for improvement under limited bandwidth conditions, while the background is transmitted complementarily without affecting the overall visual effect. Ultimately, this strategy maximizes the overall quality of the image while preserving its subjective visual integrity.

The network structure of this paper is shown in Figure 4.

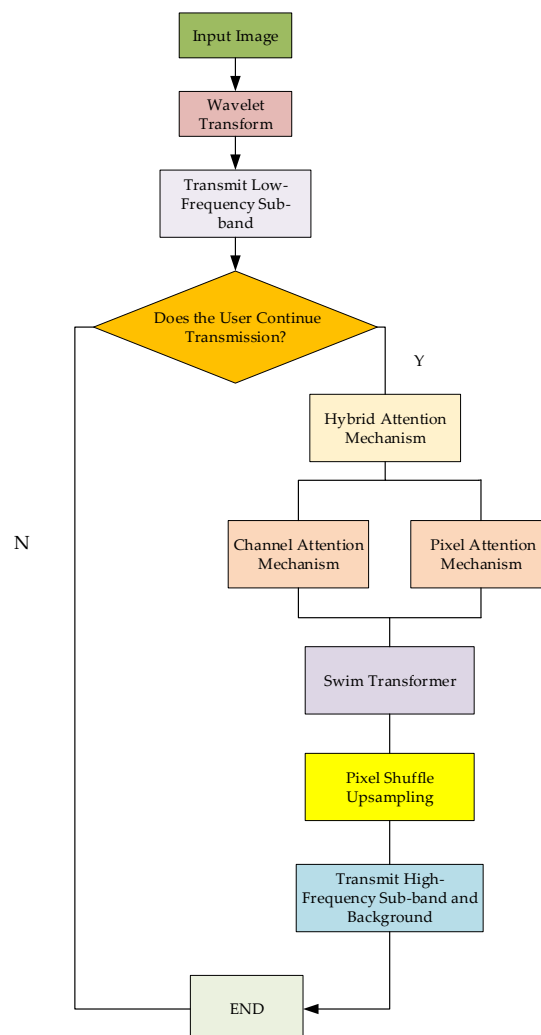


Figure 4. Network structure.

4. Experimental Results and Analysis

4.1. Data Sets and Parameter Settings

The dataset used in this experiment is composed of two parts. The first part includes images of power towers captured by UAVs in Unsignalized area in southeastern Jilin Province, China. This section contains a large volume of images, and several representative examples are selected for the experiments presented in this paper (as shown in Figure 5). The second part is the TowerDetection-v1 dataset, which comprises 2813 images, featuring numerous power line images from signal-free regions. To ensure the model's generalization and reliability, the dataset is divided into a training set, validation set, and test set, with proportions of 70%, 20%, and 10%, respectively. Specifically, the training set contains 1969 images, the validation set consists of 563 images, and the test set includes 281 images.



Figure 5. Images taken by the group in a signal-free area in Jilin.

During the data preparation stage, to enhance the model's robustness in complex backgrounds and varying environments, accelerate convergence, and improve training stability, all images were normalized to ensure pixel values fall within the range [0, 1]. Special care was taken to preserve key target areas, such as towers and wires, ensuring they remain intact and unobstructed, thereby improving detection accuracy.

The experiments were conducted on a system equipped with a RTX 3080 GPU (10 GB RAM), running Ubuntu 20.04.6. Development was carried out using CUDA 12.0, Python 3.8, and the deep learning framework PyTorch 1.8. The training process consisted of two stages: In the first stage, a batch size of eight was used, with the SGD optimizer set to an initial learning rate of 0.002 over 200 epochs. The second stage employed a cosine decay strategy for the learning rate, starting at 0.0002, with a weight decay of 0.02. The Swin Transformer module was configured with a window size of 8×8 and an encoding length of 256.

During the training process, the model is evaluated on the validation set after each epoch, primarily assessing image reconstruction quality using PSNR and SSIM metrics. Hyperparameters are adjusted based on validation results to optimize model performance. The images in the validation set differ from those in the training set in terms of shooting angle, lighting conditions, and background complexity, ensuring the model's adaptability and robustness across diverse scenes.

4.2. Evaluation Indicators

PSNR (Peak Signal-to-Noise Ratio) is a commonly used metric for evaluating the effectiveness of image compression, particularly in the field of image and video compression. It measures image quality by comparing the error between the original image and the compressed or restored image. A higher PSNR value generally indicates that the compressed image is closer in quality to the original.

PSNR is calculated as:

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

$$MSE = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n (I(i,j) - K(i,j))^2$$

Here MAX is the maximum possible pixel value of the image, and MSE is the mean squared error between the original and compressed or recovered images, calculated as the average of the squared pixel intensity differences between the two images. The image size is $m \times n$ meaning it consists of m rows and n columns; $I(i,j)$ represents the pixel value of the original image at position (i,j) ; $K(i,j)$ denotes the pixel value of the compressed or processed image at the same position. It should be noted that PSNR does not always align with human visual perception of image quality, so it is often used in conjunction with SSIM (Structural Similarity Index).

SSIM (Structural Similarity Index or Structural Similarity Index Measure) is a metric used to assess the visual similarity between two images, particularly in the fields of image compression and super-resolution. It evaluates the quality of a processed image by examining three aspects: brightness, contrast, and structure. This approach provides a closer approximation to how the human eye perceives image quality, offering a more reliable assessment than traditional metrics.

The SSIM calculation formula is:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Here x and y are the two image windows being compared; μ_x and μ_y represent the mean luminance of x and y , respectively; σ_x^2 and σ_y^2 denote the variances of x and y ; and σ_{xy} is the covariance of x and y . The constants c_1 and c_2 are used to prevent division by zero. Typically, $c_1 = (k_1L)^2$ and $c_2 = (k_2L)^2$ where L is the dynamic range of the pixel values, and $k_1 = 0.01$ and $k_2 = 0.03$.

4.3. Ablation Experiment

To thoroughly evaluate the contribution of each module in the image transfer algorithm, we incrementally add modules and conduct experiments on the dataset. All experiments are evaluated using two metrics: PSNR and SSIM, to assess the quality of image reconstruction. The experimental results are presented in Table 2.

Table 2. Results of ablation experiments with different modules stacked in sequence.

Experiment	Wavelet Transform	Hybrid Attention Mechanism	Swin Transformer Module	Pixel Shuffle Upsampling Module	PSNR/dB	SSIM
Experiment I	✓				27.032	0.813
Experiment II	✓	✓			29.913	0.854
Experiment III	✓	✓	✓		33.174	0.916
Experiment IV	✓	✓	✓	✓	40.2	0.982

As shown in Table 2, this study conducts ablation experiments by gradually adding modules to evaluate their contributions to the model's performance. In the base model (Experiment I), only the wavelet transform is used for image decomposition and reconstruction, without the hybrid attention module, Swin Transformer module, or Pixel Shuffle upsampling module. This configuration yields a PSNR of 27.032 dB and an SSIM of 0.813. However, since only low-frequency information is preserved, the reconstructed image lacks

detail, with blurred edges and significant texture loss, particularly in the high-frequency regions.

In Experiment II, the hybrid attention mechanism module is added, enhancing the feature weights of channels and pixels. This addition improves the PSNR by 2.881 dB and the SSIM by 0.041, significantly enhancing the retention of key details and improving the model's feature extraction capabilities. Experiment III further incorporates the Swin Transformer module, which captures global information through the local window self-attention mechanism. This improves the PSNR by 3.261 dB and the SSIM by 0.062, allowing the model to effectively capture contextual information while retaining local details, thereby enhancing the overall visual quality of the image.

Finally, Experiment IV adds the Pixel Shuffle upsampling module, which improves spatial resolution by rearranging the channel information of the feature map. This addition increases the PSNR by 6.031 dB and the SSIM by 0.046, further enhancing the image clarity. The Pixel Shuffle module also mitigates blurring and checkerboard artifacts often associated with traditional interpolation methods, resulting in sharper edges and texture details. The experimental results demonstrate that the modules complement each other, enhancing the overall performance of the model. The complete model achieves optimal performance in both PSNR and SSIM, verifying the necessity and effectiveness of each module.

To further verify the effectiveness of the modules, we show the feature map comparison after each ablation experiment. As shown in Figure 6, with the increase of modules, the extracted features are gradually enhanced, especially in details and edge information.

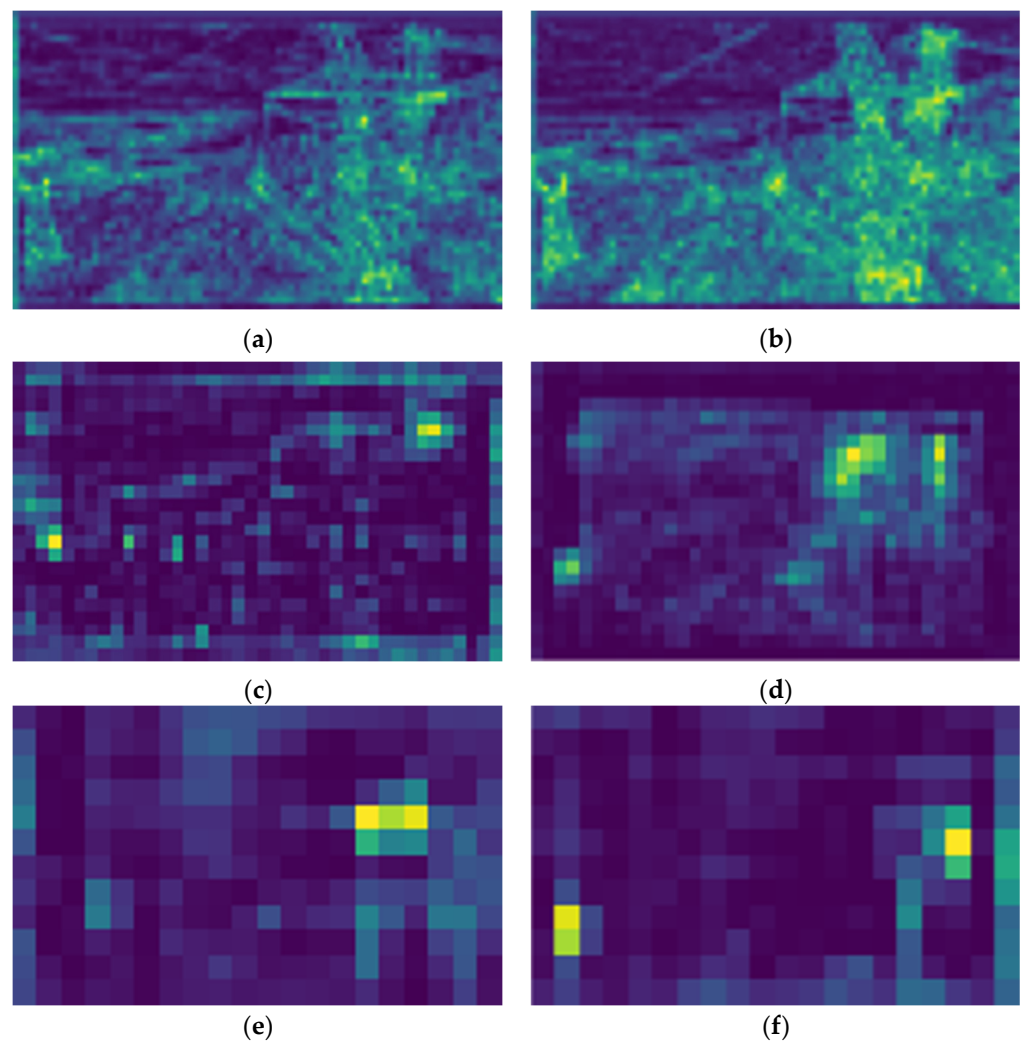


Figure 6. Comparison of feature extraction between traditional algorithm and this paper's algorithm.

As shown in Figure 6, plots (a), (c), and (e) demonstrate the feature extraction results at each stage of the traditional recovery algorithm, while plots (b), (d), and (f) show the improvement of the model's feature extraction results after the gradual addition of the hybrid attention module, the feed-forward Swin Transformer module, and the Pixel Shuffle upsampling module, respectively. In the traditional algorithm, the (a) plot has more noise, the feature distribution is dispersed, and the edges and details are not clear enough, while the (b) plot with the addition of the hybrid attention module has more concentrated features, sharper contours, and less noise, which demonstrates a higher level of feature extraction capability. (c) plot features are sparse and have fewer highlights, resulting in limited expressive power; however, the Swin Transformer-processed (d) plot has increased feature densities, especially in the key regions, with significantly enhanced highlights, capturing more localized patterns and texture information. The features of (e) plot are concentrated in localized areas with insufficient information coverage, while (f) plot after adding Pixel Shuffle up-sampling shows a wider distribution of features, and useful information is detected in several key locations, which improves the adaptability of the model in complex scenes. Overall, with the gradual addition of each module, the feature extraction effect is significantly improved, and the robustness and global sensing ability of the model are effectively enhanced.

4.4. Comparison Experiment

The transmission time in this study is measured experimentally and is defined as the total duration from the initiation of image transmission to the successful recovery of the complete image. The measurement procedure is as follows: The timer starts when the low-frequency subbands of the image begin transmission and stops when all necessary high-frequency subbands have been transmitted and the complete image is successfully recovered. This approach ensures an accurate recording of the entire process, from the start of data transmission to the end of image recovery.

To simulate a realistic transmission environment in an unlicensed region, a network emulator is used in the experiments, restricting the bandwidth to 110 KB/s and introducing a network delay of 180 ms. This setup allows for an effective evaluation of the algorithm's adaptability and efficiency in a low-bandwidth environment.

Several well-established traditional progressive image transmission methods are widely used to gradually transmit and display images under limited bandwidth conditions. The core concept of these methods is to transmit low-resolution or basic information first, followed by progressively more detailed information, thereby enhancing the image quality step by step. Although these conventional methods are mature and widely adopted, their efficiency and performance may be insufficient in signal-free regions. To address this limitation, the proposed algorithm integrates the attention mechanism with a local self-attention model to improve transmission quality under restricted bandwidth conditions.

To verify the superiority of the proposed method in unlicensed regions, comparative experiments were conducted against mainstream methods, including MPRNet [25], DiffLight [26], Uformer GAN [27], Fourier Prior Architecture [28], and Progressive Disentangling [29]. The results of these experiments are presented in Table 3.

Table 3. Comparison of this paper's method with mainstream progressive image recovery methods.

Method	PSNR (dB)	SSIM	Bandwidth Consumption (KB)	Transmission Time (s)
MPRNet	39.7	0.970	400	3.5
DiffLight	37.5	0.945	450	3.9
Uformer GAN	38.8	0.955	420	4.0
Fourier Prior Architecture	38.2	0.950	430	3.8
Progressive Disentangling	36.9	0.940	460	4.2
Methodology of this paper	40.2	0.982	300	2.8

As shown in Table 3, the proposed algorithm significantly reduces bandwidth consumption and transmission time while maintaining high image quality, outperforming mainstream methods such as MPRNet, DiffLight, and Uformer GAN. Specifically, the algorithm achieves a PSNR of 40.2 dB and an SSIM of 0.982, demonstrating superior performance in both bandwidth consumption and transmission efficiency.

The algorithm's efficiency is attributed to the use of wavelet transform and the Pixel Shuffle module, which prioritize the transmission of low-frequency subbands and employ efficient feature mapping, minimizing redundant information. As a result, the bandwidth consumption is reduced to only 300 KB, and the transmission time is shortened to 2.8 s—about 1 s less than other algorithms. Additionally, the hybrid attention and Swin Transformer modules reduce computational complexity while preserving edge and texture details, ensuring excellent performance even in extreme environments.

The experimental results confirm that the proposed method achieves a well-balanced performance in complex scenarios such as unsignaled regions, improving image recovery accuracy while significantly reducing resource overhead. Future research can explore multi-module co-optimization and real-time adaptive strategies to address more complex transmission requirements and further enhance efficiency by integrating deep learning-based compression techniques.

Overall, the proposed algorithm offers an innovative solution to image transmission challenges in signal-free regions, achieving an optimal balance between transmission efficiency and resource utilization while ensuring high-quality image recovery.

Quantitative metrics in image denoising, such as PSNR and SSIM, provide a numerical evaluation of algorithm performance but may not fully capture the differences in image quality perceived by the human eye. Therefore, the effectiveness of image restoration is also assessed through visual comparison. As shown in Figure 7, a comparison of the image recovery effect of the proposed method with that of other algorithms.

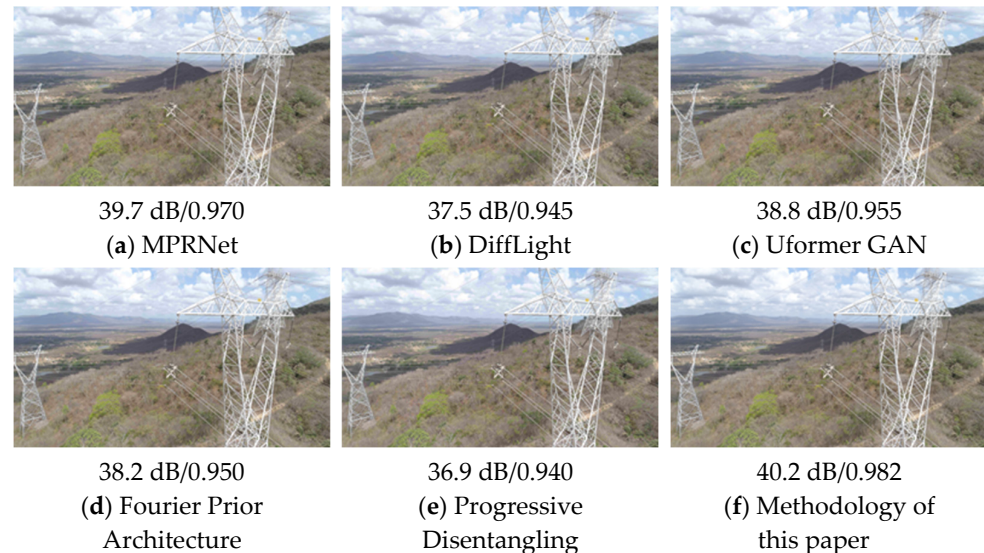


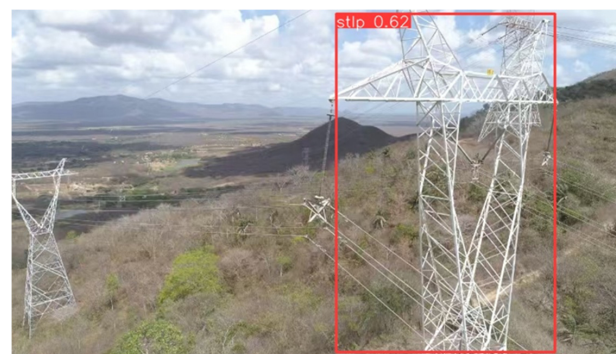
Figure 7. Visualization of image recovery by different methods.

4.5. Image Target Detection

In order to verify the effectiveness of the proposed algorithm for transmission line monitoring in signal-free areas, we conducted a two-step detection process. First, we take the images taken by the group in the signal-free region of Jilin Province and input them into the YOLOv10 model for testing in order to observe the recognition performance of the recovered images in complex environments. Second, in order to test the generalization ability of the algorithm, we used the public dataset TowerDetection-v1 as a benchmark to further evaluate the structural integrity and detection accuracy of the recovered images.

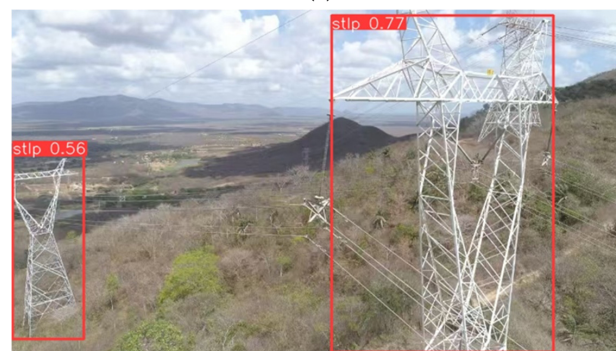
In each experiment, we use images recovered by different algorithms as inputs to analyze the performance differences of each method in the structural detection of transmission towers and lines. This process not only helps us quantify the effectiveness of different recovery strategies, but also visualizes the advantages and disadvantages of each algorithm in image detail processing.

As shown in Figure 8a, the image recovered using YOLOv10 to detect the Progressive Disentangling method only identifies the right transmission tower, with a confidence level of 0.62. In Figure 8b, the image recovered using YOLOv10 to detect the MPRNet method achieves a confidence level of 0.56 for the left transmission tower and 0.77 for the right transmission tower. In contrast, as shown in Figure 8c, the image recovered using our method achieves a confidence level of 0.74 for the left transmission tower and 0.78 for the right transmission tower. Overall, the confidence levels for both transmission towers detected using YOLOv10 with our method are higher than those of the other two methods.



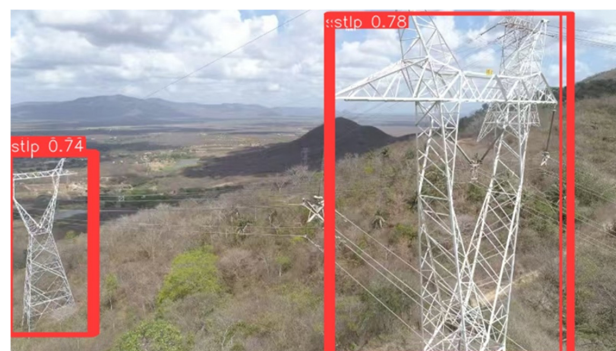
Progressive Disentangling

(a)



MPRNet

(b)



Methodology of this paper

(c)

Figure 8. The dataset of images taken by the group in the unsignalized area of Jilin Province was recovered by different methods and then inputted into YOLOv10 for detection results.

Figure 9 illustrates the varying performances of different methods on a publicly available dataset, as evaluated by YOLOv10. As shown in Figure 9a–d, the Progressive Disentangling method excels in restoring low-frequency structures but struggles with high-frequency detail processing, resulting in incomplete detection of key transmission line details. Figure 9e–h show that the MPRNet method offers strengths in edge sharpening and basic structure detection but has limited retention of complex textures, impacting detection comprehensiveness and accuracy. In contrast, Figure 9i–l demonstrate that our method significantly enhances detail recovery through the integration of a hybrid attention mechanism and the Swin Transformer. This approach accurately identifies the fine structure of transmission lines and towers, exhibiting superior robustness and generalization in detection.

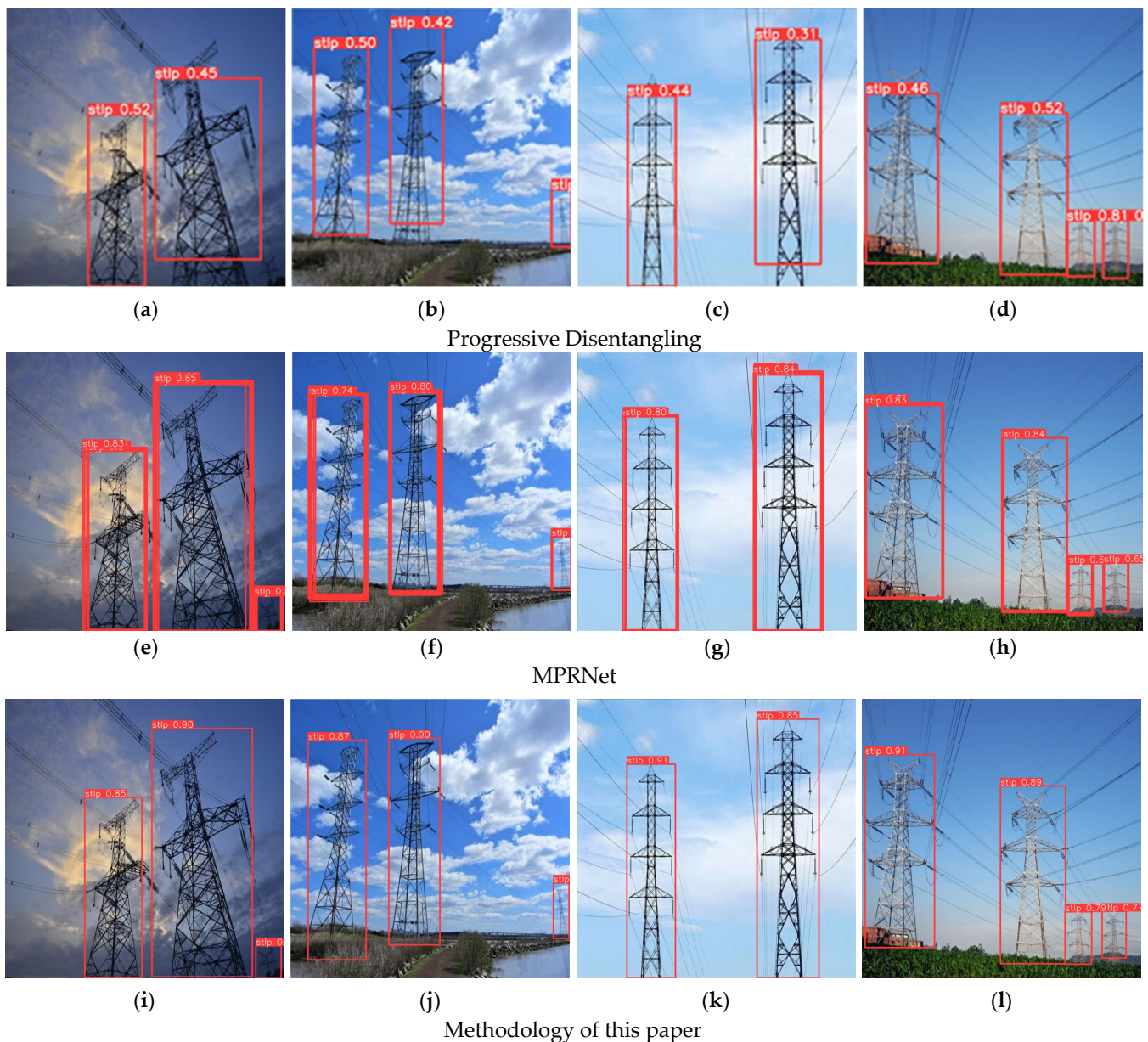


Figure 9. Visualization of image recovery by different methods. (a–d) is method Progressive Disentangling; (e–h) is method MPRNet; (i–l) is method Methodology of this paper.

The experimental results highlight that the quality of the recovered image is crucial for accurate and comprehensive target detection. Compared to other algorithms, our method performs exceptionally well on both datasets, demonstrating excellent detail

fidelity and edge detection. Additionally, it maintains stable detection results across various complex environments.

By prioritizing the transmission of critical regions and utilizing the Pixel Shuffle upsampling module to enhance resolution, our algorithm achieves efficient image transmission and recovery under limited bandwidth conditions. Whether applied to complex environments in self-constructed datasets or detection tasks in public datasets, our method demonstrates exceptional adaptability and performance, offering a reliable solution for transmission line monitoring in signal-free areas.

5. Conclusions

This study proposes a progressive image transmission method based on BeiDou short message communication and an attention mechanism to meet the needs of transmission line monitoring in signal-free areas and address the bottlenecks of traditional image transmission methods in low-bandwidth environments. By integrating wavelet transform with a hybrid attention mechanism, Swin Transformer, and Pixel Shuffle upsampling module, the proposed method significantly enhances the clarity and visual quality of image restoration while reducing transmission time and bandwidth consumption.

Experimental results show that, compared with other mainstream methods, the algorithm proposed in this paper not only achieves significant improvements in image quality (PSNR, SSIM) but also demonstrates higher transmission efficiency in terms of resource consumption. Through the incremental optimization of each sub-module, the final algorithm achieves the optimal balance between the subjective visual quality of images and resource utilization under limited bandwidth conditions. Furthermore, visualization comparison experiments and target detection verify the practical applicability of the recovered images, confirming the effectiveness and superiority of the proposed method for surveillance in signal-free areas.

However, the current method still has certain limitations. First, although the hybrid attention mechanism enhances the ability to recover fine details, the overall computational complexity remains high, potentially restricting its use on resource-constrained devices. Second, the algorithm's real-time adaptability may be insufficient in scenarios with extreme dynamic environments or frequent changes in signal conditions. Future research can further optimize multi-module collaboration strategies, explore adaptive transmission mechanisms to meet more complex scene requirements, and integrate compression techniques with deep learning to enhance the efficiency of image transmission. This study offers an innovative solution to address image transmission challenges in signal-free regions, demonstrating high application value.

Author Contributions: Conceptualization, X.J. and X.Y.; methodology, X.J.; software, X.Y.; validation, X.J., X.Y. and H.Y.; formal analysis, H.G.; investigation, Z.Y.; resources, X.J.; data curation, X.J.; writing—original draft preparation, H.Y.; writing—review and editing, X.Y.; visualization, X.Y.; supervision, X.J.; project administration, H.G.; funding acquisition, H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Technology Project of State Grid Co., Ltd. (SGJLTH00XTJS2400378).

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to confidentiality reasons related to laboratory data.

Conflicts of Interest: Author Hongliu Yang was employed by the company State Grid Jilin Electric Power Co., Ltd. Tonghua Power Supply Company. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Himeur, Y.; Boukabou, A. Robust image transmission over powerline channel with impulse noise. *Multimed. Tools Appl.* **2017**, *76*, 2813–2835. [[CrossRef](#)]
2. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. ECA-Net: Efficient channel attention for deep convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020.
3. Pattichis, C.S.; Kyriacou, E.; Voskarides, S.; Pattichis, M.S.; Istepanian, R.; Schizas, C.N. Wireless telemedicine systems: An overview. *IEEE Antennas Propag. Mag.* **2002**, *44*, 143–153. [[CrossRef](#)]
4. Zhang, Y.; Li, K.; Li, K.; Wang, L.; Zhong, B.; Fu, Y. Image super-resolution using very deep residual channel attention networks. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018.
5. Dubovik, O.; Schuster, G.L.; Xu, F.; Hu, Y.; Bösch, H.; Landgraf, J.; Li, Z. Grand challenges in satellite remote sensing. *Front. Remote Sens.* **2021**, *2*, 619818. [[CrossRef](#)]
6. Chen, L.; Zhang, H.; Xiao, J.; Nie, L.; Shao, J.; Liu, W.; Chua, T.-S. SCA-CNN: Spatial and channel-wise attention in convolutional networks for image captioning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
7. Kose, S.; Koçtak, E.; Hascicek, Y.S. An overview on the use of satellite communications for disaster management and emergency response. *Int. J. Emerg. Manag.* **2012**, *8*, 350–382. [[CrossRef](#)]
8. Zhao, H.; Kong, X.; He, J.; Qiao, Y.; Dong, C. Efficient image super-resolution using pixel attention. In *Computer Vision—ECCV 2020 Workshops: Glasgow, UK, 23–28 August 2020, Proceedings, Part III 16*; Springer International Publishing: Berlin, Germany, 2020.
9. Alam, A.; Sheeba, P. A review of automatic driving system by recognizing road signs using digital image processing. *J. Inform. Electr. Electron. Eng. (JIEEE)* **2021**, *2*, 1–9. [[CrossRef](#)]
10. Xiao, J.; Guo, H.; Yao, Y.; Zhang, S.; Zhou, J.; Jiang, Z. Multi-scale object detection with the pixel attention mechanism in a complex background. *Remote Sens.* **2022**, *14*, 3969. [[CrossRef](#)]
11. Liang, J.; Cao, J.; Sun, G.; Zhang, K.; Van Gool, L.; Timofte, R. Swinir: Image restoration using swin transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021.
12. Li, B.; Li, X.; Lu, Y.; Liu, S.; Feng, R.; Chen, Z. Hst: Hierarchical swin transformer for compressed image super-resolution. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer Nature: Cham, Switzerland, 2022.
13. Nascimento, V.; Laroca, R.; Lambert, J.d.A.; Schwartz, W.R.; Menotti, D. Combining attention module and pixel shuffle for license plate super-resolution. In Proceedings of the 2022 35th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Natal, Brazil, 24–27 October 2022; IEEE: Piscataway, NJ, USA, 2022; Volume 1.
14. Zhang, D.; Zhang, D. Wavelet transform. In *Fundamentals of Image Data Mining: Analysis, Features, Classification and Retrieval*; Springer: Berlin, Germany, 2019; pp. 35–44.
15. Munson, D.C. A note on Lena. *IEEE Trans. Image Process.* **1996**, *5*, 3. [[CrossRef](#)]
16. Islam, A.; Amin, R.; Sharan, M.H.; Ushno, J.A.; Islam, K.M.T.; Hossen, D.; Hossain, M.S.; Alam, M.S. Designing an autonomous framework to detect and classify the fabric defects using wavelet transform and neural network. In *AIP Conference Proceedings*; AIP Publishing: Melville, NY, USA, 2023; Volume 2788.
17. Hua, Q.; Chen, L.; Li, P.; Zhao, S.; Li, Y. A pixel-channel hybrid attention model for image processing. *Tsinghua Sci. Technol.* **2022**, *27*, 804–816. [[CrossRef](#)]
18. Liu, N.; Han, J.; Yang, M.-H. PiCANet: Pixel-wise contextual attention learning for accurate saliency detection. *IEEE Trans. Image Process.* **2020**, *29*, 6438–6451. [[CrossRef](#)] [[PubMed](#)]
19. Hu, J.; Li, S.; Gang, S. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
20. Islam, A.; Amin, R.; Sharan, M.H.; Ushno, J.A.; Islam, K.M.T.; Hossen, D.; Hossain, M.S.; Alam, M.S. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 11–17 October 2021.
21. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 6999–7019. [[CrossRef](#)] [[PubMed](#)]
22. Anderson, P.G. Linear pixel shuffling for image processing: An introduction. *J. Electron. Imaging* **1993**, *2*, 147–154. [[CrossRef](#)]
23. Odena, A.; Dumoulin, V.; Olah, C. Deconvolution and checkerboard artifacts. *Distill* **2016**, *1*, e3. [[CrossRef](#)]
24. Sunil, C.K.; Jaidhar, C.D.; Patil, N. Tomato plant disease classification using multilevel feature fusion with adaptive channel spatial and pixel attention mechanism. *Expert Syst. Appl.* **2023**, *228*, 120381.
25. Mehri, A.; Ardakani, P.B.; Sappa, A.D. MPRNet: Multi-path residual network for lightweight image super resolution. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Virtual, 5–9 January 2021.
26. Feng, Y.; Hou, S.; Lin, H.; Zhu, Y.; Wu, P.; Dong, W.; Sun, J.; Yan, Q.; Zhang, Y. DiffLight: Integrating Content and Detail for Low-light Image Enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024.
27. Ouyang, X.; Chen, Y.; Zhu, K.; Agam, G. Image restoration refinement with Uformer GAN. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024.

28. Nehete, H.; Monga, A.; Kaushik, P.; Kaushik, B.K. Fourier Prior-Based Two-Stage Architecture for Image Restoration. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024.
29. Liu, G.; Yue, H.; Yang, J. Efficient Light Field Image Super-Resolution via Progressive Disentangling. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024.

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.