remote sensing

MDPI

Article

Lightweight Remote Sensing Image Dehazing via Dual-Stage Feature Calibration and Hierarchical Context Aggregation

Yufeng He [1,2], Cuili Li [1,3,*], Mingdeng Shi [1,2] and Xu Li [1,2]

[1] College of Information Engineering, Tarim University, Alar 843300, China; heyufeng@taru.edu.cn (Y.H.)
[2] Key Laboratory of Tarim Oasis Agriculture, Ministry of Education, Tarim University, Alar 843300, China
[3] College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
* Correspondence: licuili@taru.edu.cn

Highlights

What are the main findings?

- We propose a lightweight remote sensing image dehazing method that integrates the information-aware feature calibration and a hierarchical context aggregation mechanism for feature representation reinforcement.
- The proposed method achieves state-of-the-art performance on multiple public datasets, outperforming 10 competing methods across all quantitative and qualitative tests.

What are the implications of the main findings?

- The method effectively restores haze-degraded remote sengsing images, significantly enhancing the visual quality of optical images under complex haze conditions.
- It provides a more reliable solution for hazy image restoration, benefiting various applications like crop classification and land cover mapping.

Abstract

Remote sensing images (RSIs) are frequently degraded by atmospheric haze, which introduces color distortion and contrast reduction, thereby impeding downstream applications. Existing models often struggle with non-uniform haze distributions, high computational costs, and the loss of local texture details. To address these challenges, this paper proposes a lightweight Feature Self-Recalibration Network (FSRNet) for efficient remote sensing image dehazing. FSRNet adopts a symmetric encoder–decoder architecture as its backbone and utilizes parameter-free pixel shuffle and unshuffle operations for multiscale feature resampling to preserve complex spatial details. The core of FSRNet lies in the specially designed Feature Self-Recalibration Module (FSRM), which consists of two key components: the Dual-Stage Feature Calibration Block (DFCB) and the Hierarchical Context Aggregation Block (HCAB). Specifically, the DFCB statistically splits features into informative and redundant parts, independently recalibrating them through a simplified channel attention mechanism to enhance representation in heterogeneous haze regions. Simultaneously, the HCAB integrates a non-local haze perception branch and a local detail enhancement branch in parallel, enabling the model to perceive global haze density while preserving fine-grained textures. Experimental results on multiple authoritative synthetic and real-world remote sensing datasets demonstrate that FSRNet achieves state-of-the-art dehazing performance. With only 0.865 M parameters and 8.622 G MACs, FSRNet strikes a superior balance between restoration quality and computational efficiency, making it highly suitable for real-time deployment on resource-constrained platforms.

Check for updates

Academic Editor: Silvia Liberata Ullo

Received: 7 January 2026
Revised: 30 January 2026
Accepted: 9 February 2026
Published: 11 February 2026

Copyright: © 2026 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.

Keywords: remote sensing image; haze removal; deep learning; image restoration; feature calibration

1. Introduction

Remote sensing image (RSI) plays a vital role in meteorological observation, disaster assessment, military reconnaissance, and land resource surveys due to its ability to capture large-scale surface information [1]. However, scattering and absorption of incident light by atmospheric aerosols and water vapor introduce radiometric attenuation, global contrast degradation, and color distortion in acquired imagery, markedly reducing its reliability [2]. Consequently, haze-contaminated images impede downstream applications like land-cover classification, change detection [3], and object recognition [4,5], which require high-fidelity details and radiometric consistency [6]. Therefore, developing efficient and accurate dehazing algorithms for remote sensing images is not only a prerequisite for improving data quality, but also an urgent requirement to ensure the reliability and robustness of subsequent high-level remote sensing applications.

Under hazy conditions, optical remote sensing imaging follows the atmospheric scattering model [7]: incident irradiance reflected by the surface is scattered and absorbed by atmospheric aerosols, and ambient light is superimposed along the propagation path, yielding the sensor-captured radiance:

$$I(x) = J(x)t(x) + A(1 - t(x)) \tag{1}$$

where $I(x)$ denotes the haze image at pixel x, $J(x)$ is the clear image to be restored, and A is the global atmospheric light. $t(x) = e^{-\beta d(x)}$ is the transmission of the atmospheric medium, exponentially related to scene depth $d(x)$ and the atmospheric scattering coefficient β. This model presents two core challenges: it is severely ill-posed (three unknowns $J(x)$, A, and $t(x)$, but only one observation $I(x)$), and non-uniform variations in $d(x)$ and β (due to RSI's large coverage) further increase ambiguity. Thus, single image dehazing requires inverting these unknowns to recover scene irradiance while suppressing non-uniform haze degradation.

Early dehazing methods fall into two categories: quality enhancement-based approaches (e.g., histogram equalization, Retinex) [8–10] treat haze as generic degradation but ignore physical mechanisms, leading to over-enhancement or distortion under dense and inhomogeneous haze. Model-based methods [11–13] add image priors to the atmospheric scattering model, offering high interpretability but failing when scenes violate prior assumptions, causing transmission estimation errors and artifacts. Recent data-driven methods [14–16] use convolutional neural networks (CNNs) or Transformers to learn a direct mapping $f : I \to \hat{J}$ from hazy–clean pairs, implicitly encoding physical haze cues via pixel-wise or perceptual loss. However, these models have key limitations: they treat features uniformly (causing under-enhancement in dense haze and over-processing in clear regions), fail to fuse global haze distribution and local details (leading to artifacts in large RSIs), and are often too bulky for real-time deployment. To address these, we propose a lightweight network that adaptively recalibrates features by haze density and hierarchically fuses global-local information.

We propose FSRNet, a feature self-recalibration network with a symmetric encoder–decoder backbone. It uses pixel shuffle and unshuffle operations for parameter-free multiscale resampling and inserts a Feature Self-Recalibration Module (FSRM) at each scale. FSRM integrates two core blocks: Dual-Stage Feature Calibration Block (DFCB), which splits features into informative/redundant parts and recalibrates them with lightweight

attention, and Hierarchical Context Aggregation Block (HCAB), which parallelly captures global haze density and local details. Using 1 × 1 and depth-separable convolutions, FSRNet delivers strong performance with low computational cost.

The primary contributions of this paper can be succinctly outlined as follows:

- We propose FSRNet, a feature self-recalibration network that effectively recovers clear RSIs with only 0.865M model parameters and lightweight computation.
- We design the Dual-Stage Feature Calibration Block (DFCB), which splits features by information content and recalibrates them to enhance heterogeneous haze representation.
- We introduce the Hierarchical Context Aggregation Block (HCAB), which parallelly captures global haze density and local details via a dual-granularity mechanism.
- Extensive experiments show FSRNet achieves state-of-the-art performance on synthetic and real RSI datasets, with downstream task tests validating its effectiveness.

2. Materials and Methods

2.1. Related Work

This section reviews representative dehazing techniques based on image enhancement, physical priors, and deep learning. It then summarizes the latest progress in lightweight-oriented image dehazing and analyzes the advantages and limitations of these methods.

2.1.1. Image-Enhancement-Based Dehazing Methods

Image enhancement approaches improve visibility by directly manipulating contrast, edges, or frequency components without physical modeling. Classical techniques include global histogram equalization (HE) and its adaptive variant CLAHE [9], though the latter remains prone to blocking artifacts. Retinex-based methods [8,17] estimate illumination to recover reflectance but often fail under dense haze due to inaccurate estimation. Furthermore, frequency-domain filters like wavelet transform [18] and homomorphic filtering [19], alongside gamma correction [10], color adjustment [20], and white balance [21], are frequently employed. While effective for thin haze, these methods ignore the underlying haze physics, frequently leading to detail loss, color distortion, and poor performance in spatially non-uniform haze conditions.

2.1.2. Physics-Model-Based Dehazing Methods

Physics-based dehazing methods build upon the atmospheric scattering model and constrain the ill-posed problem with handcrafted priors. The dark channel prior (DCP) [11] is the most influential, enabling effective transmission estimation but causing over-enhancement and color distortion in sky or bright regions. Subsequent works have refined DCP with homomorphic filtering [22], light–dark priors with adaptive patch strategies [23], or efficient local extremum priors [24] for better robustness and efficiency. Beyond transmission refinement, several studies relax the globally uniform atmospheric-light assumption with spatially varying priors [25–27]. Others replace handcrafted constraints with statistical or global regularities, such as the color attenuation prior [12], color ellipsoid fitting [28], or haze-line priors [13], while iterative virtual-depth strategies [29] further enhance estimation accuracy. These physics-based methods have strong interpretability and perform well in typical scenes, but they rely on specific priors and fail in scenes with large bright regions (e.g., water, snow, white buildings), causing inaccurate transmission and atmospheric light estimation and limiting robustness in complex RSI scenes.

2.1.3. Deep-Learning-Based Dehazing Methods

Data-driven dehazing has become mainstream, learning end-to-end hazy-to-clear mappings without handcrafted priors. Early works such as AOD-Net [30] and LightDehazeNet [31] reformulated the atmospheric scattering model into end-to-end CNNs to directly restore clear images. Subsequent studies extend this paradigm with illumination compensation [15], asymmetric U-Nets with attention refinement [14], and depth-assisted joint optimization [32], achieving SOTA results on synthetic and real benchmarks. Purely data-driven methods abandon the atmospheric scattering model entirely: GridDehazeNet [33], GCANet [34], and INP-Net [35] boost haze-density perception and texture restoration via multi-scale fusion, attention and implicit neural prompting, while MixDehazeNet [16] integrates multi-scale convolution and attention to handle uneven haze distributions effectively. RSI-specific dehazing methods have also emerged: Li and Chen [36] proposed a coarse-to-fine two-stage framework, while Wen et al. [37] designed minimal encoder–decoder architectures with multi-view interaction for efficiency-accuracy balance in large-scale high-resolution RSIs. Despite these advances, deep learning-based methods still face key challenges: over-reliance on synthetic training data, high computation for large-scale images, and poor performance on non-uniform haze in complex RSI scenes.

2.1.4. Recent Advances in Efficient and Lightweight Dehazing

Recent lightweight dehazing studies focus on balancing computational efficiency, parameter scale and dehazing performance. LFD-Net [38] uses a reformulated ASM with multi-scale convolutions and gated fusion-attention for real-time dehazing and improved downstream object detection. Progressive optimization networks split dehazing into low-cost initial restoration and iterative refinement: PFONet [39] uses multi-stream feature extraction and feedback for fast inference, while PSR-Net [40] designs a spatially adaptive feature block to handle RSI diversity. RI-SCNN [41], the first spiking dehazing network, mimics retinal ON/OFF pathways, reducing energy consumption to 1.072 mJ per image, and LDFormer [42] optimizes transformers with selective feedforward and cross-layer attention for efficient feature extraction. He et al. [43] design LRSDN with axial depthwise convolution and hybrid attention, achieving superior performance with <0.1 M parameters. For general image restoration, Zhong et al. [44] and Nguyen et al. [45] proposed TUNet and UFR-GAN, achieving excellent multi-task performance with lightweight designs. These works show that lightweight dehazing now relies on structure simplification, mechanism innovation and domain adaptation, yet still faces challenges in handling dense haze and extreme weather conditions.

2.2. Methodology

This section elaborates on the proposed Feature Self-Recalibration Network (FSRNet) for RSI dehazing. We first describe its symmetric encoder–decoder architecture, which uses pixel shuffle/unshuffle for multi-scale feature resampling. Its core innovations are two key blocks: the Dual-Stage Feature Calibration Block (DFCB), which decomposes and recalibrates features by information richness, and the Hierarchical Context Aggregation Block (HCAB), which fuses local details and global haze density to boost feature representation. Finally, we introduce the loss function for network training.

2.2.1. Overall Architecture

To address non-uniform haze and restore RSI details, we design FSRNet, a lightweight symmetric encoder–decoder network. As shown in Figure 1a, its architecture has three stages: a shallow convolutional layer for feature embedding, mapping 3-channel RGB images to a high-dimensional latent space; a concise multi-scale encoder that extracts

features and reduces resolution via cascaded convolution and downsampling, which cuts memory usage, expands the receptive field of subsequent convolutions, and incurs no extra computation; an encoder-symmetric decoder with cascaded convolution and upsampling to process extracted features and output haze-free images. Following existing empirical practices [46,47], a Tanh activation layer is used as the final layer to generate restored images.

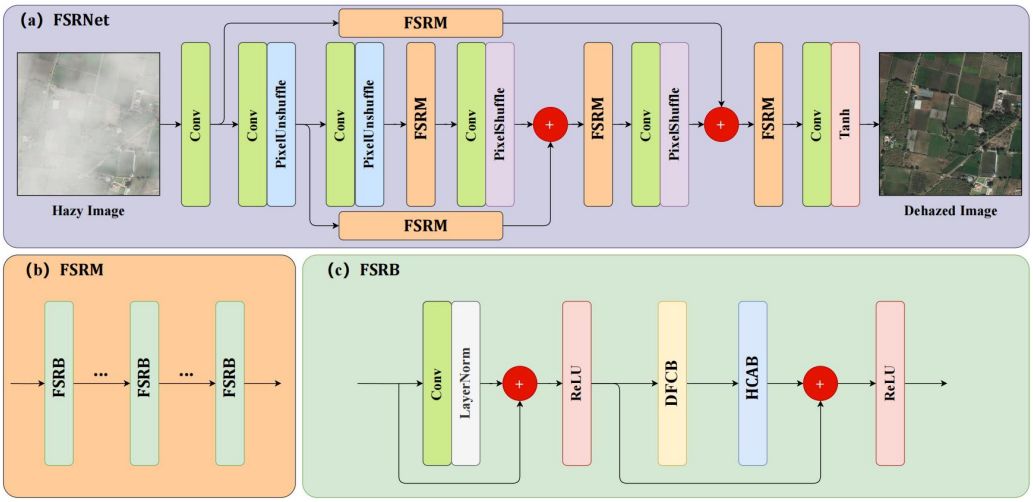


Figure 1. The overall structure of the proposed FSRNet. (**a**) is the schematic structure of the FSRNet based on the symmetric encoder and decoder. (**b**) The composition of FSRM, the core module of FSRNet, which consists of several cascaded FSRBs. (**c**) The internal architecture of the FSRB, which has a fundamental residual structure and two key components, DFCB and HCAB.

The FSRM (highlighted in the orange box in Figure 1a is embedded into the network as the core innovation. An FSRM is integrated at each decoder feature scale to boost scale-specific representational capacity. Skip connections are established between encoder and decoder features at the same spatial scale [48], allowing encoder multi-scale features to be directly fused into decoder features via residual connections, and FSRMs are also applied to these skip connections for feature enhancement. These designs jointly strengthen shallow–deep layer information exchange and facilitate gradient backpropagation during training. FSRM (Figure 1b) comprises three cascaded Feature Self-Recalibration Blocks (FSRBs), whose detailed structure is shown in Figure 1c. FSRB is built on a classic residual structure, with core components DFCB and HCAB (detailed in Sections 2.2.2 and 2.2.3). DFCB splits input features by information richness, recalibrates each partition independently via attention, and fuses the refined features; HCAB enhances high- and low-frequency feature representation via dual-granularity aggregation.

FSRNet uses pixel shuffle [49] for feature upsampling, instead of traditional transposed convolution or linear interpolation. Originally designed for image super-resolution, pixel shuffle realizes upsampling via feature map channel grouping and spatial rearrangement. As shown in Figure 2, given a low-resolution feature (H_i, W_i, ch_i) with $ch_i = C \times r^2$ (C: group number, r: upsampling factor), pixel shuffle splits the feature into C channel-wise groups; each group rearranges spatial pixels to generate an r-times upsampled single-channel feature, yielding an output of $(H_i \times r, W_i \times r, C)$. Notably, pixel shuffle has no learnable parameters and relies only on rearrangement, ensuring high computational efficiency, and it also avoids checkerboard artifacts and blurred edges induced by transposed convolution and linear interpolation. These advantages are critical for RSI dehazing, where preserving complex ground object details is essential. For downsampling, we adopt pixel unshuffle, the inverse operation of pixel shuffle, which reduces spatial resolution while preserving details via channel expansion, mitigating information loss relative to traditional pooling.

We quantitatively validate the effectiveness of pixel shuffle and unshuffle for RSI dehazing via ablation experiments with and without these operations.

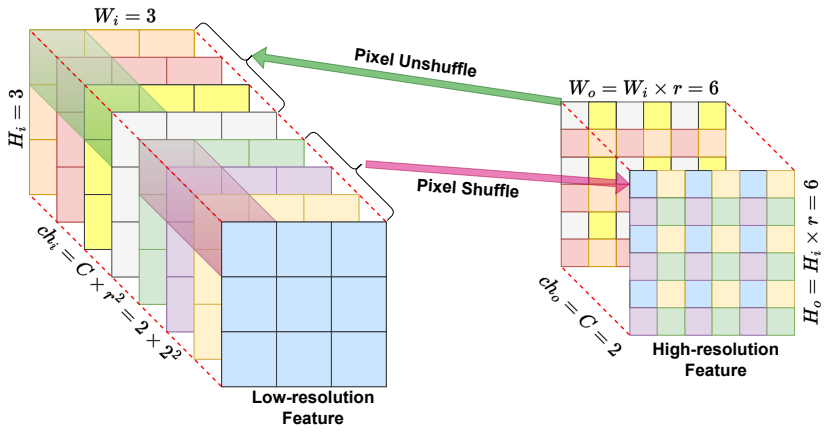


Figure 2. Pixel shuffle and unshuffle operations. Different colors represent different channels of the feature map.

2.2.2. Dual-Stage Feature Calibration Block

Haze in RSIs exhibits spatial heterogeneity due to large imaging coverage, leading to uneven degradation across image sub-regions. Standard convolution treats all feature regions equally, causing insufficient processing of dense haze areas and computational waste from over-processing clear regions. Additionally, traditional convolutions generate substantial spatial and channel redundancy in feature extraction [50]. To address these issues, we propose the Dual-Stage Feature Calibration Block (DFCB) inspired by [50,51], which splits feature calibration into two stages (Figure 3): Information-aware Feature Splitting, which partitions features via input statistics and threshold gating, and Attention-based Feature Reconstruction, which independently enhances split features with simplified channel attention [52] and fuses them via element-wise addition.

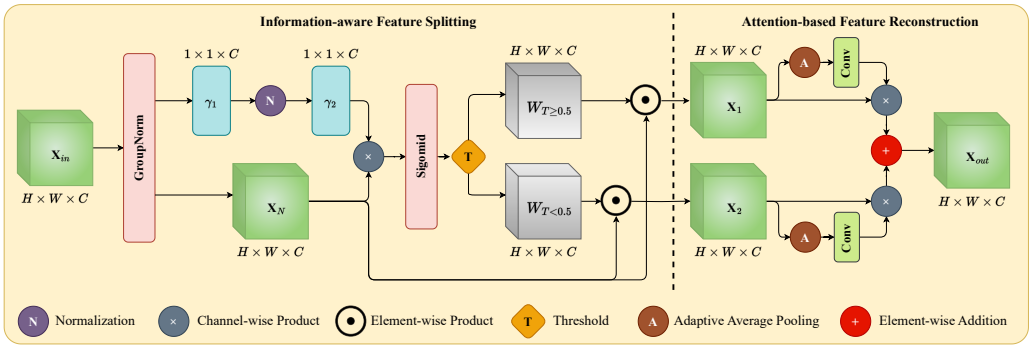


Figure 3. The structure of the proposed DFCB.

In the Information-aware Feature Splitting stage (left of the vertical dashed line in Figure 3), we use the learnable scaling factors of group normalization to estimate feature informativeness. For an input feature map $X_{in} \in \mathbb{R}^{H \times W \times C}$ (C: channel number, H/W: spatial resolution), group normalization is first performed as:

$$X_N = GN(X_{in}) = \gamma_1 \frac{X_{in} - \mu}{\sqrt{\sigma^2 + \varepsilon}} + \beta \tag{2}$$

where $GN(X_{in})$ and X_N denote group normalization operation for the input feature map and the normalized feature, respectively. σ and μ are the standard deviation and mean of X_{in}. γ_1 and β are channel-independent learnable scaling and shifting parameters of the

normalization, and a small constant ε ensures numerical stability in Equation (2). $\gamma_1 \in \mathbb{R}^C$ is not an input feature statistic, but a channel-level feature importance weight learned from the overall input sample distribution during training. We thus use γ_1 to distinguish feature informativeness and normalize it to γ_2 as:

$$\gamma_{2,i} = \frac{\gamma_{1,i}}{\sum \gamma_1}, i = 1, 2, \cdots, C \tag{3}$$

After reweighting by γ_2, X_{GN} is mapped to (0, 1) via the sigmoid function to generate W, a metric of feature informative potential. Subsequently, W is decomposed into two parts ($W_{T \geq 0.5}$ and $W_{T<0.5}$) by threshold gating, and each part is multiplied with the normalized feature to yield two same-shaped sparse feature maps X_1 and X_2, corresponding to informative and redundant features, respectively.

Following [50], redundant features contain non-trivial uncertain information (arising from similar but non-identical feature extraction patterns), so they are not discarded but enhanced independently with X_1 in the Attention-based Feature Reconstruction stage (right of the vertical dashed line in Figure 3). We adopt a simplified channel attention [52] with adaptive average pooling and 1×1 convolution to model channel-wise weights, performing separate enhancement for X_1 and X_2 to match the human visual system's differential perception of information-rich and -sparse regions. The two enhanced branches are then summed to generate the final feature map, with stage 2 formulated as:

$$X_{out} = X_1 \cdot Conv_1(P_a(X_1)) + X_2 \cdot Conv_1(P_a(X_2)) \tag{4}$$

where $Conv_1(\cdot)$ and $P_a(\cdot)$ means the 1×1 convolution and adaptive average pooling, respectively.

DFCB uses no high-computation components, with only two 1×1 convolutions introduced in its second stage, enabling lightweight feature representation enhancement with minimal parameter and computation overhead. Unlike DFCB's information-aware split and independent recalibration, typical attention-based enhancement modules (RCAB [53], CBAM [54]) perform global equal treatment of feature maps and enhance features via spatial and channel dimensions. We quantitatively compare the effectiveness of these modules for RSI dehazing in the ablation study (Section 4.1.2).

2.2.3. Hierarchical Context Aggregation Block

Conventional convolution operations, limited by fixed receptive fields, exhibit distinct patterns in information extraction. Shallow convolutional layers capture high-frequency features (e.g., local edges and textures), while deep layers focus on low-frequency information (e.g., global illumination and structures). Both are indispensable for remote sensing image dehazing: global information effectively models the spatial distribution of non-uniform haze, and local information aids in recovering ground object details to avoid edge blurring in dehazed results. While existing studies often use skip connections between shallow and deep layers for global-local information fusion, this overly simplistic design cannot fully exploit contextual cues. Inspired by the human visual system, which integrates holistic perception and focal attention in parallel, we propose the Hierarchical Context Aggregation Block (HCAB). It adopts a dual-granularity strategy to aggregate global structures and local details, built on a residual structure with two parallel branches: one for global haze density perception and the other for local detail enhancement.

As illustrated in Figure 4, the input feature X_{in} first passes through a Group Normalization (GN) layer and a 1×1 convolutional layer for channel expansion. It is then split into two equal-sized feature tensors X_1 and X_2. X_1 is fed into the non-local haze perception branch (noted as NHPB in Figure 4) to model global haze distribution, and X_2

passes through the local detail enhancement branch (LDEB) to refine high-frequency details. Subsequently, the outputs X_G (from NHPB) and X_L (from LDEB) are concatenated along the channel dimension, and then processed by a 1×1 convolutional layer for channel compression and feature fusion, generating the aggregated feature X_A. Finally, X_A is summed with normalized input features X_N in line with residual learning principles to produce the output of the HCAB. The overall workflow can be formally expressed as follows:

$$X_N = GN(X_{in}) \tag{5}$$

$$X_1,\, X_2 = Split(Conv_1(X_N)) \tag{6}$$

$$X_G = \mathbf{B}_{NHP}(X_1) \tag{7}$$

$$X_L = \mathbf{B}_{LDE}(X_2) \tag{8}$$

$$X_A = Conv_1(Concat(X_G,\, X_L)) \tag{9}$$

$$X_{out} = X_A + X_N \tag{10}$$

Here, $Split(\cdot)$ and $Concat(\cdot)$ denote feature channel-wise split and concatenation operations, respectively. $\mathbf{B}_{NHP}(\cdot)$ and $\mathbf{B}_{LDE}(\cdot)$ represent the two parallel branches of the HCAB, and their detailed architectures will be elaborated in subsequent paragraphs. All other symbols in the formula adhere to the conventions established earlier.

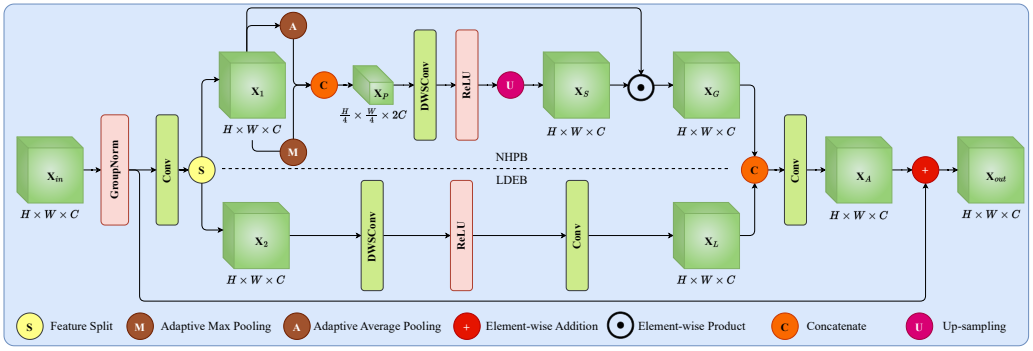


Figure 4. The structure of the HCAB.

In the non-local haze perception branch (NHPB), the input feature X_1 first undergoes the adaptive average pooling and the adaptive max pooling with a downsampling factor of 4, respectively. The downsampled results are then concatenated along the channel dimension to form X_P, followed by a depthwise separable convolutional layer (written as DWSConv in Figure 4) and a ReLU activation for nonlinear mapping. The processed feature is subsequently upsampled to the original spatial resolution of X_1, yielding X_S, which serves as the global haze distribution weights. Finally, X_S is multiplied with X_1 to generate the global haze-aware feature X_G. The NHPB operation can be expressed mathematically as follows:

$$X_P = Concat(P_a(X_1),\, P_m(X_1)) \tag{11}$$

$$X_S = Up(ReLU(DWSConv(X_P))) \tag{12}$$

$$X_G = X_1 \odot X_S \tag{13}$$

where $P_a(\cdot)$ and $P_m(\cdot)$ denote adaptive average pooling and adaptive max pooling, respectively. DWSConv represents a depthwise separable convolution layer, which consists of a 3×3 depthwise convolution together with a 1×1 pointwise convolution, $\odot$ signifies the Hadamard product (element-wise multiplication), and $Up(\cdot)$ denotes the upsampling operation, which we implement via nearest-neighbor interpolation in this work. In contrast,

the local detail enhancement branch (LDEB) has a more lightweight structure, and its computational flow can be expressed as

$$X_L = Conv_1(ReLU(DWSConv(X_2)). \tag{14}$$

It sequentially applies a depthwise separable convolutional layer, a ReLU activation, and a 1×1 convolutional layer to enhance the module's representational capacity with only a small number of additional parameters. It refines local details of the input feature X_2 on its original spatial scale, ensuring the restored image retains finer textures and edges.

2.2.4. Loss Function

In the model training phase, we adopt L2 loss as the optimization objective for the network, which can be expressed as follows:

$$\mathcal{L}_2 = \|\widehat{J(x)} - J(x)\|^2 \tag{15}$$

where $J(x)$ represents the corresponding ground truth image and $\widehat{J(x)}$ denotes the restored image. Compared to L1 loss, the L2 loss imposes stronger gradient penalties on large errors through its quadratic term, compelling the network to prioritize correcting severely distorted regions, e.g., ground objects covered with dense haze. This approach can effectively address the non-uniform haze distribution, which is the primary challenge in remote sensing image dehazing, alleviate local over-enhancement (e.g., artifacts at the edges of images) caused by the L1 loss, and ultimately yield visually superior dehazing results. Although existing studies frequently employ hybrid loss functions, such as various linear combinations of L1, L2, SSIM, and perceptual loss, their weight settings lack theoretical guidance, requiring manual design and extensive trial-and-error. Furthermore, these combinations risk introducing optimization conflicts, such as the inconsistency between the SSIM loss and pixel-level loss functions.

3. Results

3.1. Implementation Details

The number of cascaded FSRBs within each FSRM module is set to 2. The model is trained using the L2 loss function with a batch size of 24 for 100 epochs. We employ the Adam optimizer with a cosine annealing scheduler, decreasing the initial learning rate from 1×10^{-3} to 1×10^{-7}. Mixed precision training is applied to accelerate computation and optimize memory usage. Training samples are shuffled in each epoch, 90% of the them are randomly cropped and the remaining 10% are resized to generate 256×256 image patches. To enhance model generalization, random flipping and rotation augmentations are also applied to training data. The entire training pipeline is implemented using Python 3.11 and PyTorch 2.1.0 with an NVIDIA GeForce RTX 3090 GPU.

3.2. Compared Methods

FSRNet is compared against representative methods, including physical model-based approaches: DCP [11], IDeRs [29], and SMIDCP [22]; and deep learning-based models: AOD-Net [30], GCANet [34], AUNet [14], LDNet [31], CASM [15], MDNet [16], and the transformer-based MABDT [55]. All methods use official implementations and are evaluated under the same hardware and software configurations to ensure a fair comparison.

3.3. Experiments on Paired Remote Sensing Datasets

We conducted extensive experiments on multiple remote sensing haze datasets, which provide paired hazy/haze-free images under predefined atmospheric parameters for accu-

rate, objective evaluation. RSHD [43] is a high-resolution (0.05–0.3 m GSD) remote sensing haze dataset with 42,240 training and 3060 test pairs (1024 × 1024 px), synthesized via a Perlin-noise-based atmospheric scattering model under thin, moderate, and dense haze conditions with both homogeneous and inhomogeneous distributions. SateHaze1k [56] has 1200 triplets of hazy/haze-free RGB and SAR images from GF-2 and GF-3 satellites, we only used optical pairs, evenly split into thin, moderate, and thick haze subsets (512 × 512 px). RSID [57] has 1000 haze/haze-free pairs (256 × 256 px) under dense and non-homogeneous haze conditions in military scenes, with a 9:1 training/testing split. RICE [58] is a real-world remote sensing haze dataset collected from Google Earth via the atmospheric veil layer visibility toggling, containing 500 pairs haze/haze-free images with a resolution of 512 × 512 pixels. RSHD and SateHaze1k are divided into subsets by haze concentration, while RICE and RSID do not. In addition, RSID, RSHD, and SateHaze1k simulate non-uniform haze.

To quantify performance, we employ five standard full-reference metrics: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Learned Perceptual Image Patch Similarity (LPIPS) [59], Feature Similarity Index (FSIM) [60], and CIEDE2000. These metrics evaluate dehazing quality from pixel-wise accuracy, structural consistency, perceptual similarity, feature-level similarity, and color fidelity, respectively. For PSNR, FSIM, and SSIM, higher values are preferred, while for LPIPS and CIEDE2000, lower values indicate superior restoration.

3.3.1. Haze Removal on RSHD Dataset

The RSHD test set comprises six subsets each containing 510 samples, categorized by haze distribution (uniform/non-uniform) and density (thin/moderate/thick). Figure 5 presents visual results of representative samples, with rows for haze conditions and columns for methods; the first column is hazy inputs and the last is ground truth (GT) haze-free images. The visual comparison shows that traditional prior-based methods (DCP, IDeRs) have obvious color distortion and unnatural brightness, especially under non-uniform dense haze. SMIDCP removes haze effectively but suffers from severe over-enhancement and inaccurate color restoration. CNN-based methods (AOD-Net, GCANet, CASM) reduce haze but often produce oversmoothed edges or residual haze. LDNet and AUNet restore more natural colors but occasionally introduce over-enhancement artifacts. MDNet achieves competitive visual quality but has slight local color bias. MABDT performs well on light haze but fails to fully remove dense haze. In contrast, FSRNet restores clearer edges, more accurate colors, and balanced brightness, effectively eliminating haze residues across all levels to produce results closest to GT. Quantitatively (Table 1), FSRNet achieves the best scores in PSNR, SSIM, FSIM, and CIEDE2000, and ranks second-best in LPIPS. These results demonstrate that FSRNet effectively balances structural fidelity, color consistency, and perceptual quality, outperforming both classical and state-of-the-art learning-based approaches on the high-resolution RSHD dataset.

Table 1. Full reference assessment scores of comparison methods dehazing on RSHD dataset. ↑ means a better performance with the higher value, while ↓ means the opposite case. The best evaluation values are marked in bold.

Method	DCP	MABDT	IDeRs	SMIDCP	AOD-Net	GCANet	AUNet	LDNet	CASM	MDNet	FSRNet
LPIPS ↓	0.312	**0.200**	0.411	0.331	0.266	0.214	0.233	0.235	0.300	0.215	0.202
PSNR ↑	18.736	28.590	14.714	18.392	20.473	27.239	26.221	23.548	22.987	27.962	**30.873**
SSIM ↑	0.796	0.879	0.634	0.760	0.830	0.873	0.872	0.854	0.817	0.853	**0.888**
FSIM ↑	0.880	0.972	0.802	0.846	0.869	0.963	0.963	0.909	0.891	0.970	**0.983**
CIEDE ↓	15.187	5.130	21.433	14.866	12.189	5.912	7.117	8.937	9.860	5.447	**3.816**

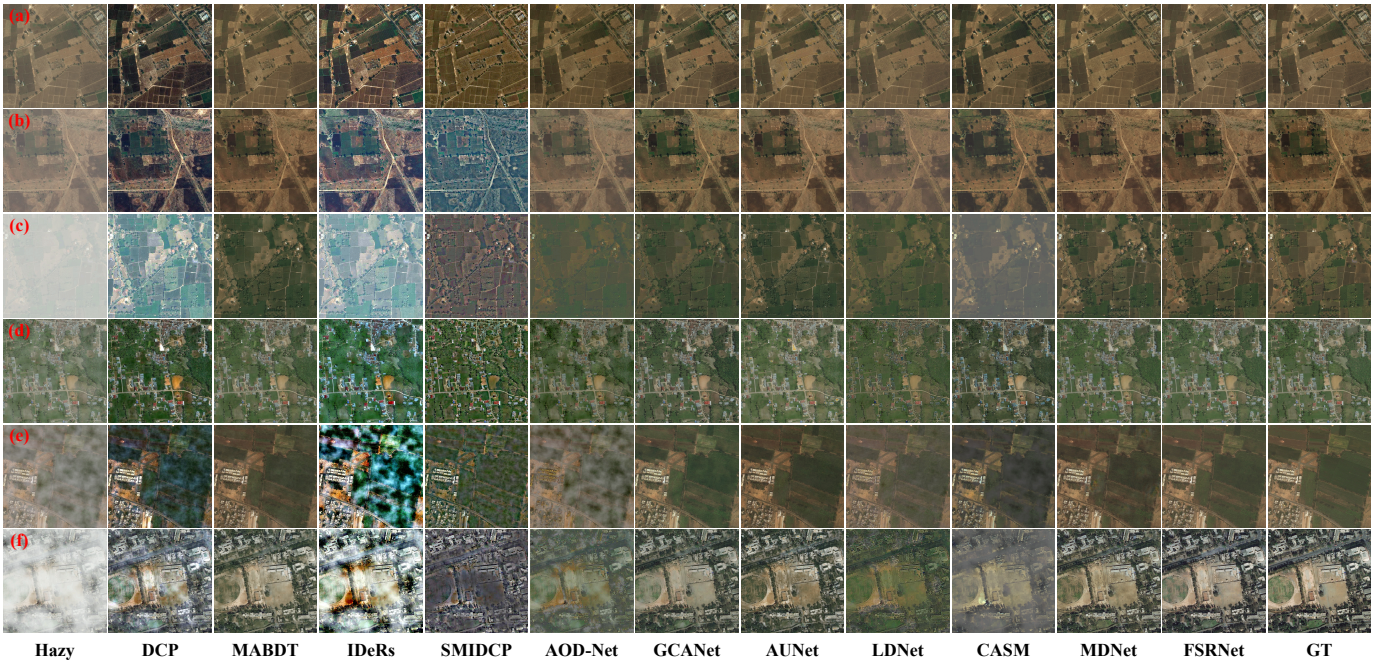


Figure 5. Visual comparison of dehazed results on RSHD dataset with various haze conditions: (**a**) homogeneous thin haze, (**b**) homogeneous moderate haze, (**c**) homogeneous dense haze, (**d**) inhomogeneous thin haze, (**e**) inhomogeneous moderate haze, and (**f**) inhomogeneous dense haze. "Hazy" indicates the input hazy image, and "GT" shows the reference image without haze.

3.3.2. Haze Removal on SateHaze1k Dataset

The SateHaze1k test set is categorized into "Thick", "Moderate", and "Thin" subsets. As shown in Figure 6, prior-based methods (DCP, IDeRs) have limited dehazing capability, leaving residual haze (e.g., lower-left/upper-right regions in Figure 6b,c) especially under dense haze. SMIDCP removes haze well but causes severe over-enhancement (oversharpened edges, unnatural colors). Deep learning-based methods perform better overall but have limitations: AOD-Net and LDNet cannot fully remove local haze (Figure 6c), and AUNet has pixel recovery errors in Figure 6b; GCANet, CASM and MDNet lack natural color fidelity under dense haze; MABDT performs well on moderate haze but has color flaws on light/dense haze samples. In contrast, FSRNet yields the most visually appealing results with clearer structures, accurate colors and minimal haze residues, showing robustness across haze densities.

Quantitative results in Table 2 reveals that prior-based DCP performs poorly (low PSNR, high perceptual distortion), lacking generalization to complex haze. IDeRs and SMIDCP also lag behind learning-based methods in structural and perceptual fidelity. CNN-based methods show clear advantages: AOD-Net and LDNet outperform traditional methods in PSNR and perceptual error, proving learning-based strategies model haze distributions effectively. GCANet, AUNet, CASM and MDNet further improve reconstruction quality with higher PSNR and lower perceptual distortion. Notably, FSRNet consistently outperforms comparison methods in most scenarios. While MABDT shows a slight advantage on the Moderate subset, FSRNet achieves the highest PSNR, SSIM, FSIM, and lowest CIEDE scores on the Thick and Thin subsets. These results validate FSRNet's robustness in maintaining structural integrity and perceptual naturalness across diverse haze concentrations.

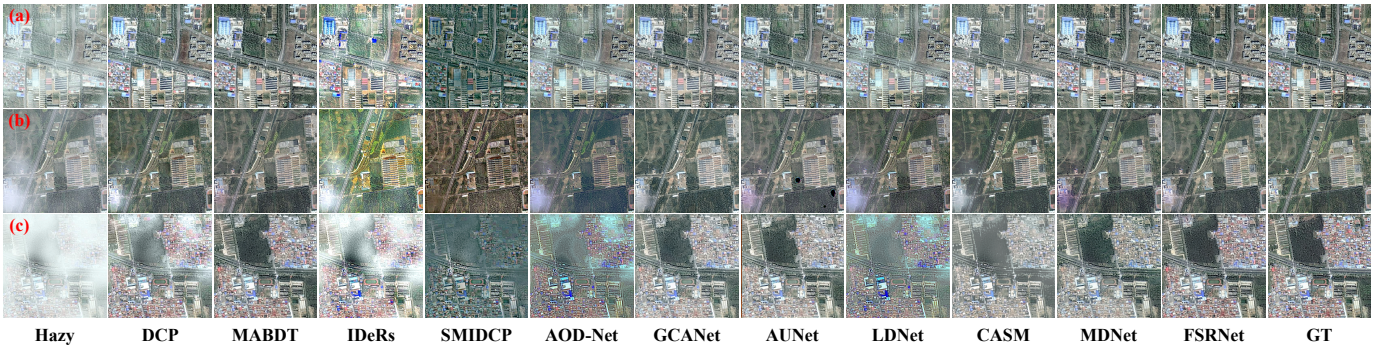


Figure 6. Visual comparison of dehazed results on SateHaze1k dataset. (**a**–**c**) represent samples with different concentrations of haze.

Table 2. Quantitative comparison of dehazing performance on SateHaze1k.

Dataset	Metrics	DCP	MABDT	IDeRs	SMIDCP	AOD-Net	GCANet	AUNet	LDNet	CASM	MDNet	FSRNet
Thick	LPIPS ↓	0.159	0.130	0.311	0.220	0.238	0.164	0.190	0.232	0.264	0.193	**0.129**
	PSNR ↑	16.377	23.772	11.754	14.956	16.521	21.059	19.909	16.506	18.269	21.899	**24.737**
	SSIM ↑	0.802	0.856	0.702	0.757	0.774	0.839	0.837	0.777	0.724	0.837	**0.868**
	FSIM ↑	0.934	0.972	0.872	0.882	0.864	0.959	0.948	0.862	0.874	0.958	**0.977**
	CIEDE ↓	14.036	6.321	21.649	16.611	15.167	8.804	9.140	15.173	12.437	7.967	**5.550**
Moderate	LPIPS ↓	0.080	**0.055**	0.321	0.153	0.175	0.087	0.104	0.116	0.154	0.079	0.069
	PSNR ↑	21.034	**27.808**	14.763	15.990	20.078	24.675	24.327	20.970	22.500	26.322	26.570
	SSIM ↑	0.931	**0.948**	0.785	0.863	0.906	0.929	0.929	0.921	0.879	0.941	0.942
	FSIM ↑	0.963	0.982	0.899	0.912	0.917	0.967	0.956	0.936	0.927	0.979	**0.981**
	CIEDE ↓	8.580	**4.804**	16.990	16.807	11.118	6.063	6.274	10.069	8.389	6.058	5.305
Thin	LPIPS ↓	0.091	**0.050**	0.278	0.134	0.098	0.077	0.071	0.095	0.107	0.085	0.052
	PSNR ↑	19.817	27.096	15.022	13.639	18.671	22.856	23.017	18.648	19.835	24.145	**27.653**
	SSIM ↑	0.888	0.926	0.771	0.833	0.870	0.909	0.906	0.873	0.874	0.913	**0.929**
	FSIM ↑	0.964	0.989	0.910	0.941	0.943	0.978	0.976	0.947	0.952	0.981	**0.991**
	CIEDE ↓	10.091	4.289	16.330	21.458	11.253	7.163	7.117	11.323	10.399	6.093	**4.009**

3.3.3. Haze Removal on RSID Dataset

On the military-oriented RSID dataset, visual results (Figure 7) show that prior-based methods (DCP, IDeRs, SMIDCP) and AOD-Net fail to effectively remove haze, leaving significant residues and blurred structures. LDNet and CASM suffer from color distortion, while GCANet and AUNet introduce unnatural textures, particularly in harbor scenes (Figure 7b). In contrast, MDNet, MABDT, and FSRNet produce images with natural colors, sharp edges, and well-preserved textures, significantly enhancing visual quality. Quantitative results in Table 3 further validate these observations. Traditional prior-based approaches yield the lowest PSNR and highest perceptual errors (LPIPS), highlighting their limitations in complex military scenarios. While subsequent learning-based methods like AUNet and MDNet show marked improvements in reconstruction accuracy and color fidelity, FSRNet consistently outperforms all competing models across all evaluation metrics. Specifically, FSRNet achieves the highest PSNR, FSIM, and SSIM, alongside the lowest LPIPS and CIEDE scores, confirming its state-of-the-art capability in restoring both structural integrity and perceptual naturalness.

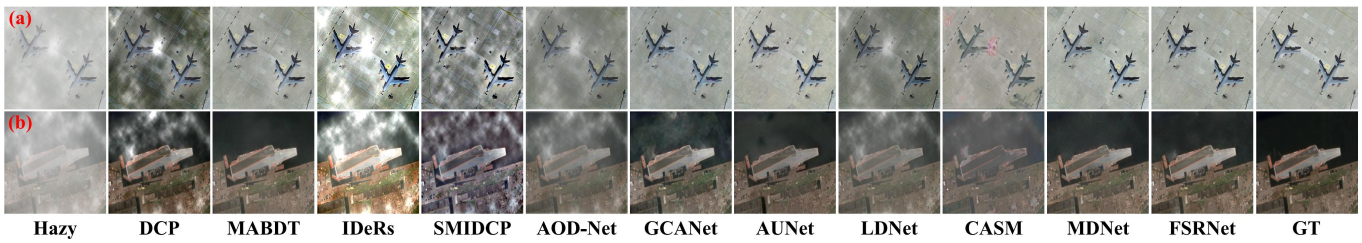


Figure 7. Visual comparison of dehazed results on RSID dataset.

Table 3. Quantitative comparison of dehazing performance on RSID dataset.

Metrics	DCP	MABDT	IDeRs	SMIDCP	AOD-Net	GCANet	AUNet	LDNet	CASM	MDNet	FSRNet
LPIPS ↓	0.160	0.042	0.277	0.199	0.144	0.083	0.054	0.125	0.203	0.048	**0.040**
PSNR ↑	17.584	24.435	14.254	16.920	19.052	22.975	25.457	19.026	20.181	25.574	**26.145**
SSIM ↑	0.880	0.962	0.713	0.838	0.901	0.938	0.957	0.912	0.879	0.959	**0.965**
FSIM ↑	0.901	0.970	0.777	0.869	0.911	0.950	0.962	0.925	0.882	0.965	**0.971**
CIEDE ↓	15.372	7.085	19.786	17.872	13.880	8.661	6.432	14.217	12.369	6.148	**5.995**

3.3.4. Haze Removal on RICE Dataset

RICE is collected from Google Earth, serving as a challenging benchmark for verifying the dehazing effectiveness on real-world RSIs. As illustrated in Figure 8, conventional prior-based methods (DCP, IDeRs, SMIDCP) show obvious over-processing. While they boost visibility and contrast of hazy inputs, their restored images deviate greatly from haze-free references, causing structural distortion and unnatural colors. Learning-based models such as AOD-Net, AUNet, LDNet, and CASM leave evident residual haze in complex scenes (e.g., Figure 8b). Conversely, MABDT, MDNet, and FSRNet achieve superior fidelity, preserving global luminance and local textures while yielding reconstructions closest to the ground truth. The comparison shows that traditional and some learning-based methods struggle with complex atmospheric conditions in real RSI data, while FSRNet attains SOTA performance with visually convincing, perceptually faithful dehazing results.

Quantitative results for RICE are listed in Table 4. Traditional prior-based methods (DCP, SMIDCP, IDeRs) perform poorly across all metrics, with high LPIPS/CIEDE and low PSNR/SSIM. This shows they cannot effectively restore fine details and color fidelity. Learning-based methods (AOD-Net, GCANet) show significant improvements; AUNet and GCANet in particular attain low LPIPS and high SSIM, indicating superior perceptual quality. Advanced deep learning models (MABDT, MDNet) also yield competitive results. However, FSRNet attains the best overall performance across nearly all metrics. It has the lowest LPIPS, and highest PSNR and SSIM. It also has the lowest CIEDE, confirming the most faithful color restoration. These results collectively confirm FSRNet not only preserves structural and perceptual details but also restores colors more accurately than existing SOTA methods, achieving the most balanced and superior dehazing performance.

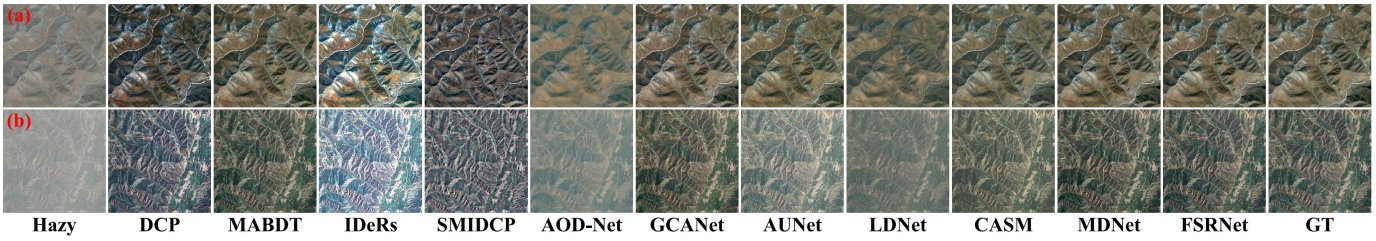


Figure 8. Visual comparison of dehazed results on RICE dataset.

Table 4. Quantitative comparison of dehazing performance on RICE dataset.

Metrics	DCP	MABDT	IDeRs	SMIDCP	AOD-Net	GCANet	AUNet	LDNet	CASM	MDNet	FSRNet
LPIPS ↓	0.231	0.073	0.369	0.337	0.274	0.077	0.075	0.183	0.157	0.092	**0.070**
PSNR ↑	18.156	31.630	15.250	16.647	20.784	29.880	28.704	23.108	24.467	31.317	**32.937**
SSIM ↑	0.832	0.951	0.612	0.718	0.834	0.946	0.946	0.873	0.922	0.949	**0.954**
FSIM ↑	0.892	0.800	0.748	0.775	0.856	0.977	0.980	0.887	0.936	**0.987**	**0.987**
CIEDE ↓	19.095	3.747	18.002	16.692	10.724	4.429	5.505	8.765	7.119	3.969	**3.632**

3.4. Experiments on Real-World Remote Sensing Images

Following full-reference evaluation on synthetic datasets, we further conduct experiments on real-world remote sensing hazy datasets, where ground-truth haze-free images are not available. In such cases, non-reference image quality assessment (NR-IQA) metrics are adopted to provide an objective evaluation of dehazing performance. These real hazy datasets present more complex and diverse atmospheric conditions compared to synthetic ones, which makes the evaluation more challenging and closer to practical applications.

We construct a Real-world Remote Sensing hazy image Dataset (RRSD100), which consists of 100 hazy remote sensing images. Among them, 62 samples with the resolution of 600×600 are manually selected from the AID [61] dataset, while the remaining 38 images with the resolution of 800×800 are collected from the DIOR [62] dataset. These samples are multi-source data collected from Google Earth by various remote imaging sensors. They cover 40 semantic scenes with pixel resolutions ranging from 0.5 to 30 m. The acquisition times, seasons, and imaging angles of the images are diverse, and the samples cover over 80 countries. Based on visual characteristics of haze imagery, such as blurred textures, low color saturation, and dim edges, the experts in remote sensing image processing manually selected these images from the AID and DIOR datasets through visual observation and the dark channel prior [11]. Most of these images exhibit inhomogeneous haze with thin or moderate density. We have made the RRSD100 dataset publicly available in Data Availability Statement.

The dehazed results of each algorithm for samples of six typical scenes in the RRSD100 dataset are shown in Figure 9. The performance of physics-based methods is generally unstable across different scenes. For instance, the IDeRs method produces apparent over-enhancement in church and storage tank scenes (samples (a) and (e)), while SMIDCP exhibits insufficient color recovery and excessive sharpening. Similarly, other competing approaches often suffer from either unnatural color restoration or blurred textures. In scenarios dominated by large green regions (e.g., farmland in Figure 9b) or scenes with rich colors (e.g., residential area in Figure 9f), most algorithms fail to recover natural and vivid colors. In contrast, the proposed FSRNet consistently generates haze-free images with visually pleasing quality across all types of scenes, which demonstrates its strong robustness and generalization ability.

Due to the absence of ground-truth haze-free counterparts in real-world remote sensing datasets, conventional full-reference metrics such as PSNR or SSIM are inapplicable. Consequently, we adopt No-Reference Image Quality Assessment (NR-IQA) methods to objectively evaluate the visual quality of dehazed images. In this study, we utilize four representative NR-IQA metrics specifically designed for haze assessment and dehazing quality: HDMHA [63], DHQI [64], RRPD [65], and VDA-DQA [66]. These metrics evaluate diverse aspects of restoration, including residual haze detection, structure preservation, distortion, and overall perceptual quality, thereby forming a multifaceted evaluation framework. To synthesize these complementary perspectives into a unified benchmark, we propose an integrated evaluation index termed "Score". Let x_i^m denote the raw assessment value of algorithm m using IQA metric i, and x_i represent the vector of scores for all compared algorithms under this metric. We first apply min-max normalization to mitigate scale discrepancies across different metrics:

$$N(x_i^m) = \frac{x_i^m - \min(x_i)}{\max(x_i) - \min(x_i)} \tag{16}$$

where $N(x_i^m) \in [0, 1]$ is the normalized value. For metrics where a lower value indicates superior quality (e.g., HDMHA), we invert the normalized scores as follows:

$$N(x_i^m) = 1 - N(x_i^m). \tag{17}$$

Consequently, for all metrics, a higher normalized value consistently reflects better dehazing performance. The final integrated score is calculated as the weighted sum of the normalized results:

$$Score = \sum_{i=1}^{n} N(x_i) \cdot w_i \tag{18}$$

where w_i is the weight of metric i, determined via the Entropy Weight Method (EWM) to reflect the discriminatory power of each IQA algorithm. Thus, "Score" serves as a unified and fair criterion, providing a comprehensive benchmark for evaluating dehazing performance on real-world remote sensing imagery.

Table 5 lists the no-reference evaluation results on RRSD100. FSRNet and MDNet obtain the best overall Scores (0.773 and 0.778, respectively), reflecting a strong balance between dehazing effectiveness and visual quality. For individual metrics, SMIDCP and LDNet reach the lowest HDMHA values. LDNet also leads in DHQI, indicating superior subjective quality. GCANet achieves the lowest RRPD, showing advantages in texture preservation, while IDeRs obtains the highest VDA-DQA, demonstrating strong detail restoration. Overall, FSRNet and MDNet offer the best balanced performance, whereas SMIDCP, LDNet, GCANet, and IDeRs excel in specific metrics, reflecting their respective focuses in haze removal.

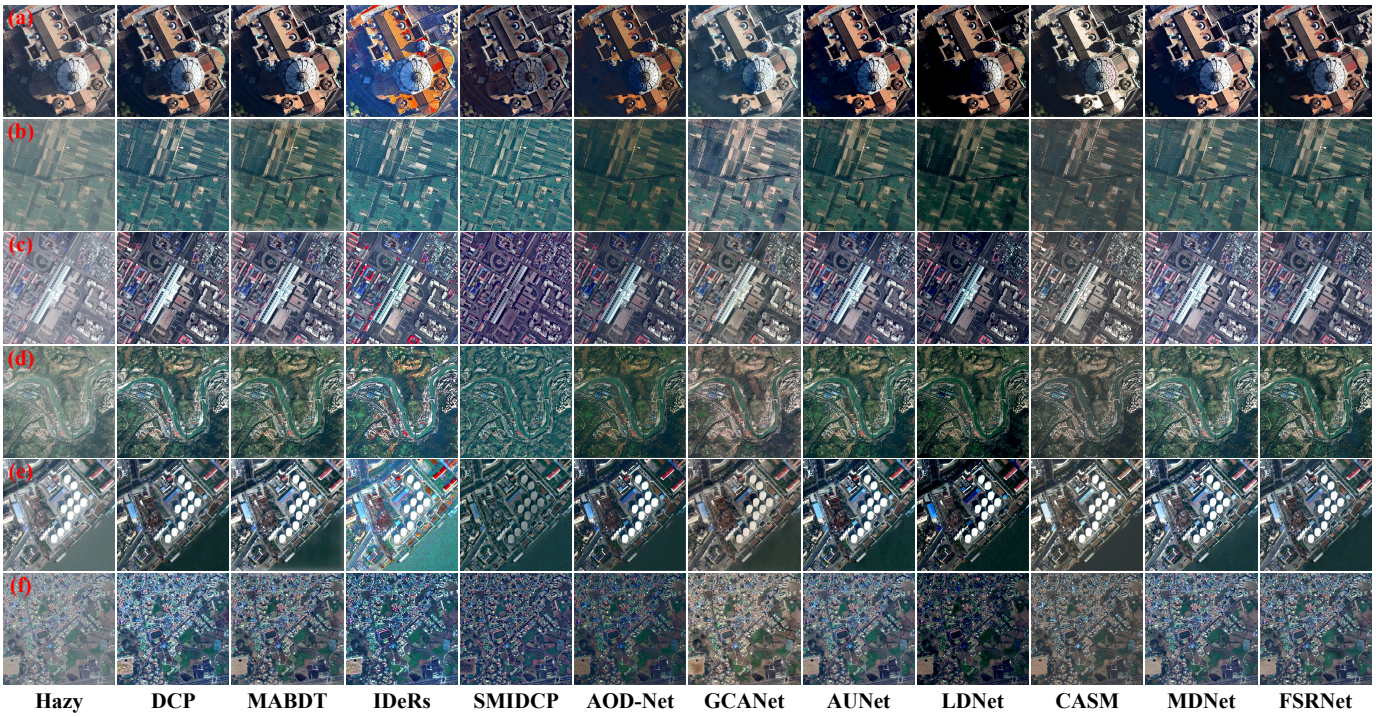


Figure 9. Visual comparison of dehazed results on real-world remote sensing images of diverse scenes: (**a**) church, (**b**) farmland, (**c**) railway station, (**d**) river, (**e**) storage tank, (**f**) residential area.

Table 5. Non-reference assessment scores of comparison methods dehazing on RRSD100 dataset. ↑ means the better performance with the higher value, while ↓ means the opposite case. Bold, underlined, and italicized values indicate the first, second, and third optimal results, respectively.

Metrics	DCP	MABDT	IDeRs	SMIDCP	AOD-Net	GCANet	AUNet	LDNet	CASM	MDNet	FSRNet
HDMHA ↓	0.014	0.032	*0.008*	**0.002**	0.003	0.096	0.025	**0.002**	0.039	0.023	0.018
DHQI ↑	64.412	64.867	53.783	65.407	68.416	64.572	67.028	**73.878**	65.288	*69.087*	69.900
RRPD ↓	25.929	26.913	38.299	35.409	35.723	**16.635**	32.342	60.955	39.562	19.644	*22.751*
VDA-DQA ↑	63.760	63.089	**66.419**	62.637	60.929	65.326	58.113	59.622	*63.969*	63.723	63.486
Score ↑	*0.711*	0.646	0.642	0.654	0.617	0.644	0.462	0.494	0.602	**0.778**	0.773

4. Discussions

This section presents a comprehensive analysis of the key findings to fully validate the effectiveness of the proposed FSRNet. Specifically, we first conduct a systematic ablation study to quantify the individual contributions of core components, thereby justifying the rationality and necessity of the architectural design. Second, we evaluate the computational efficiency and parameter complexity of FSRNet against state-of-the-art methods to verify its feasibility for deployment in resource-constrained scenarios. Subsequently, we explore the interpretability of FSRNet by visualizing feature maps to reveal potential connections with physical models. Furthermore, we examine the impact of dehazing on downstream tasks, highlighting the practical utility of FSRNet as a pre-processing step for remote sensing image understanding. Finally, we discuss the limitations of the current model and outline directions for future optimization.

4.1. Ablation Study

4.1.1. Ablation Analysis of Each Module of the FSRNet

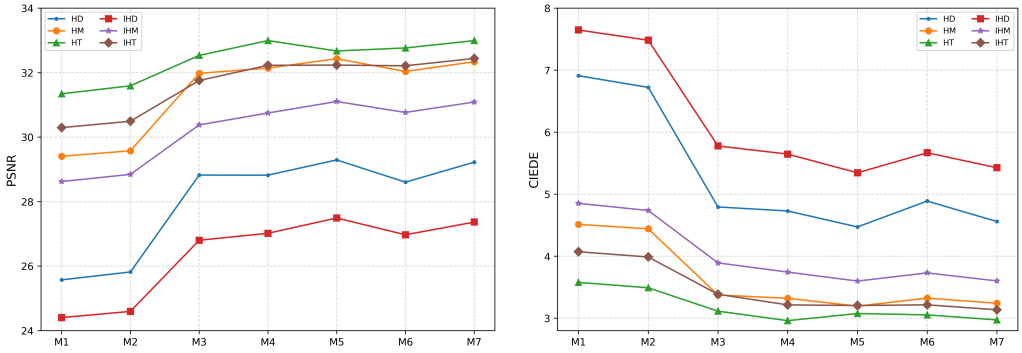
To thoroughly investigate the contribution of each component in the proposed FSRNet, we conducted a comprehensive ablation study by constructing a series of variant models. First, we constructed a baseline model, denoted as **M1**, whose overall structure is similar to FSRNet, but removes the specially designed DFCB and HCAB modules, retaining only the symmetric encoder–decoder framework, the widely used residual block, and using bilinear interpolation and max pooling for resampling, respectively. Next, we replaced the resampling layers in M1 with pixel shuffle/unshuffle operations to obtain **M2**. On this basis, we further substituted the standard residual block with the proposed DFCB (Section 2.2.2) and HCAB (Section 2.2.3), resulting in **M3** and **M4**, respectively. Finally, combining both DFCB and HCAB yields **M5**, which corresponds to the complete FSRNet. In addition, to analyze the influence of the hyper-parameter that controls the number of FSRB within each FSRM module, we constructed **M6** and **M7** by setting this number to 1 and 3, respectively.

We evaluated all seven variants (M1–M7) on the RSHD dataset, and the quantitative results are summarized in Table 6, where MACs (multiply-accumulate operations) and FLOPs (floating-point operations) are calculated by the widely used library thop (https://github.com/Lyken17/pytorch-OpCounter (accessed on 15 November 2025)) to measure the algorithm's computational overhead. The results demonstrate that pixel shuffle, DFCB, and HCAB consistently improve the dehazing performance compared to the baseline. Notably, DFCB and HCAB yield significant gains of 1.89 dB and 2.17 dB in PSNR, respectively, and their joint integration (M5) achieves the second best overall performance. Furthermore, comparing M5, M6, and M7 reveals that increasing the number of FSRBs per FSRM enhances performance, especially under heavy and non-uniform haze conditions. However, deeper stacking also results in a considerable increase in parameters, computational cost, and inference latency. When three FSRBs are employed in each FSRM (M7), the performance

improvement becomes marginal relative to the computational overhead. Therefore, to balance model complexity and dehazing effects, the final FSRNet adopts two FSRBs per FSRM. The evaluation across six subsets of RSHD (shown in Figure 10), further validates our design. In particular, DFCB and HCAB substantially enhance image quality under challenging haze conditions, such as dense and spatially non-uniform haze. These results confirm the effectiveness of the proposed architectural innovations.

Table 6. Model complexity and dehazing performance of several variants of FSRNet for ablation study. "PS" means the pixel shuffle/unshuffle technique, and "FSRBs" indicates the number of FSRB structures in each FSRM of the model.

Model Name	Network Architecture				Model Complexity				Performance	
	PS	DFCB	HCAB	FSRBs	Params (M)	MACs (G)	FLOPs (G)	Runtime (ms)	PSNR ↑	CIEDE ↓
M1	×	×	×	2	0.528	6.611	13.366	4.200	28.276	5.263
M2	✓	×	×	2	0.727	8.008	16.130	4.460	28.488	5.144
M3	✓	✓	×	2	0.527	4.882	9.688	11.230	30.380	4.057
M4	✓	×	✓	2	0.806	8.565	17.134	18.240	30.659	3.937
M5	✓	✓	✓	2	0.865	8.622	17.172	22.520	30.873	3.816
M6	✓	✓	✓	1	0.538	5.141	10.248	13.425	30.561	3.981
M7	✓	✓	✓	3	1.191	12.102	24.096	40.276	30.909	3.824



(**a**) PSNR ↑ (**b**) CIEDE2000 ↓

Figure 10. Dehazing performance of FSRNet variants ("M1"–"M7", see Table 6) on RSHD sub-test sets with diverse haze conditions (H/IH: Homogeneous/Inhomogeneous; D/M/T: Dense/Moderate/Thin; e.g., "IHD" denotes test set with inhomogeneous dense haze).

To complement the quantitative findings and intuitively validate the effects of key modules, we present qualitative comparisons of ablation variants in Figure 11. The top row displays dehazed results (red boxes highlight critical details), while the bottom row shows error maps (darker blue indicates smaller deviation from ground truth). Visual analysis confirms incremental improvements aligned with Table 6: M1 (baseline) exhibits blurred details, whereas M2 (with pixel shuffle) enhances contour sharpness. M3 (incorporating DFCB) strengthens dense haze removal and restores obscured textures, reflected in darker error maps. M4 (with HCAB) further refines color consistency, yielding natural roof color reproduction. The full FSRNet (M5) delivers the most faithful results, with complete haze removal, high-fidelity details, and minimal color deviation, with best PSNR (29.45 dB) and CIEDE (3.70) values. These observations collectively validate the synergistic benefits of each module, confirming the design rationale of FSRNet.

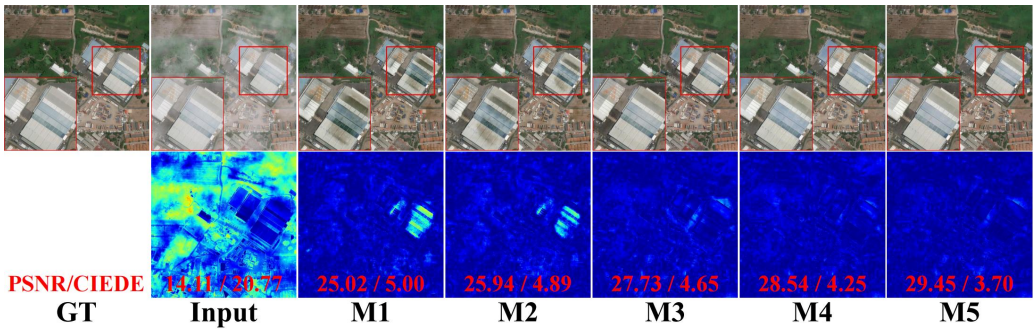


Figure 11. Qualitative comparison of variant models. The pseudo-color images depict the differences between output images and the ground truth. The red numbers respectively represent the PSNR and $CIE_{\Delta}2000$ values of the corresponding result images.

4.1.2. Ablation Comparison of Feature Calibration Modules: DFCB vs. RCAB/CBAM

In Section 2.2.2, we described the differences between the proposed DFCB module and two typical feature calibration techniques (RCAB [53] and CBAM [54]). Here, we adopt a quantitative evaluation to assess their effectiveness for remote sensing image dehazing in a more objective and impartial manner. Specifically, we take the aforementioned variant model **M3** as the baseline and replace the DFCB module of M3 with RCAB and CBAM, respectively, to construct two new variants, denoted as **M8** and **M9**. The average dehazing performance of these three models on the RSHD dataset is presented in Table 7, where "FCM" stands for the feature calibration modules used in the model, and the meanings of other notations are consistent with those in Table 6. It can be observed that compared with CBAM, DFCB achieves a significant improvement in dehazing performance with only a moderate increase in parameters and computational complexity. In contrast, although RCAB exhibits the lowest inference latency, it is not advantageous over DFCB in terms of both model scale and performance.

Table 7. Comparison of different feature calibration modules.

Model	FCM	Params (M)	FLOPs (G)	Runtime (ms)	PSNR ↑	CIEDE ↓
M3	DFCB	0.527	9.688	11.230	30.380	4.057
M8	RCAB	0.737	16.074	7.269	29.399	4.611
M9	CBAM	0.222	3.386	57.747	29.058	4.853

4.1.3. Ablation Analysis of Threshold Value T in DFCB

To address the concern about the rationale for the fixed threshold $T = 0.5$ in DFCB's threshold gating, we conducted an ablation study to evaluate the impact of different T values on model performance and feature splitting effectiveness. We adjusted the threshold T of DFCB in M3 (defined in Section 4.1.1, FSRNet variant with only DFCB) to 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. We evaluated dehazing quality (PSNR, CIEDE) on the RSHD dataset, noting that T only affects feature splitting (no change to network structure, parameters, or computational cost).

The setting of $T = 0.5$ is rooted in DFCB's work pipline: input features are normalized to [0, 1] via group normalization and sigmoid activation with the median value roughly around 0.5. This threshold ensures balanced feature splitting, preventing over-concentration in one branch and preserving meaningful information in both feature streams. Extreme values (e.g., T = 0, 1) would disable one branch entirely, breaking the dual-stage calibration mechanism, while $T = 0.5$ maintains the module's functional integrity for effective feature calibration. As shown in Table 8, the model achieves the highest PSNR and

lowest CIEDE values on both haze subsets when $T = 0.5$, confirming optimal dehazing quality and color consistency. However, when $T = 0.1$ or $T = 0.9$, performance drops sharply due to imbalanced feature splitting.

Table 8. Quantitative results of DFCB with different threshold T. "-H" and "-IH", respectively, represent the homogeneous and inhomogeneous haze test sets of the RSHD dataset.

Model	RSHD-H		RSHD-IH	
	PSNR ↑	CIEDE ↓	PSNR ↑	CIEDE ↓
M3 (T = 0.1)	30.206	4.125	28.878	4.976
M3 (T = 0.3)	30.832	3.896	29.360	4.455
M3 (T = 0.5)	31.113	3.762	29.647	4.353
M3 (T = 0.7)	30.759	3.913	29.226	4.502
M3 (T = 0.9)	30.137	4.120	28.904	4.913

4.1.4. Ablation Study of Structure Design in HCAB

In this subsection, we quantitatively analyzed some alternative designs for the HCAB module, including the downsampling factor in NHPB and the branch fusion strategy of HCAB. We conducted a systematic ablation study using M4 (refer to Table 6) as the baseline. And we evaluated the impact of varying the downsampling factor and branch fusion mechanism on both computational efficiency and dehazing performance on the RSHD dataset, with results summarized in Table 9.

Table 9. Ablation results of HCAB structure design. "DS" denotes downsampling factor and "Fusion" indicates branch fusion strategy.

Model	Architecture		Complexity		Performance	
	DS	Fusion	Params (M)	FLOPs (G)	PSNR ↑	CIEDE ↓
M4	4	concat	0.806	17.134	30.659	3.937
M4-D2	2	concat	0.806	17.330	30.674	3.898
M4-D8	8	concat	0.806	17.084	30.360	4.063
M4-F1	4	Add	0.778	16.430	30.648	3.945
M4-F2	4	NHPB-only	0.749	10.166	29.232	4.416
M4-F3	4	LDEB-only	0.708	15.662	29.725	4.253

Experimental results ("M4", "M4-D2", and "M4-D8") demonstrate that a downsampling factor of 2 yields a marginally higher PSNR but the highest FLOPs, leading to substantial computational overhead. A factor of 8 results in the lowest FLOPs but degraded performance due to excessive downsampling that loses fine-grained details of global haze distribution. In contrast, the factor of 4 achieves the optimal balance between dehazing performance and computational complexity. This configuration avoids computational overload for high-resolution remote sensing imagery while preserving sufficient global context to model non-uniform haze, thus validating its effectiveness for large-scale remote sensing scenarios. Comparisons of different fusion strategies ("M4", "M4-F1", "M4-F2", and "M4-F3") reveal that the concatenation and element-wise addition fusion have almost equal dehazing performance in both PSNR and CIEDE. Single-branch structures ("NHPB-only" and "LDEB-only") exhibit significant performance degradation. These findings indicate that the parallel concatenation structure of NHPB and LDEB optimally leverages their complementary strengths in global haze modeling and local detail recovery.

4.2. Computational Efficiency

To further evaluate the computational efficiency of different algorithms, we conducted execution latency tests under a unified experimental environment. All methods were executed on the same personal computer equipped with an Intel I7-10700 CPU, 64 GB RAM, a ZHITAI TiPro7000 2 TB SSD, and Windows 11 operating system. The source code of all competing algorithms are released at their official code repositories. Algorithms based on priors (DCP [11], CEP [28], IDeRs [29], SMIDCP [22], HazeLine [13], EVPM [24], and CAP [12]) are implemented on the Matlab R2023b platform, while deep learning methods are implemented using Python 3.11 and PyTorch 2.1.0, with parallel acceleration performed via an NVIDIA GeForce RTX 3090 GPU. We measured the average runtime of each algorithm on 100 hazy images with the resolution of 256×256, as summarized in Table 10. For deep learning models, we also reported the number of parameters and computational complexity in terms of floating-point operations (FLOPs). The results show that traditional prior-based methods lack a structured computational pipeline, making them difficult to parallelize efficiently, which leads to extensive time consumption. For instance, the runtime of HazeLine is even as high as 1607 ms for one hazy image restoration. In contrast, deep learning-based methods exhibit inference time closely related to their model complexity. The proposed FSRNet achieves a parameter size of only 0.865 M and demonstrates real-time performance under the given test environment, highlighting its potential for deployment on resource-constrained edge devices.

Table 10. Computational efficiency and model complexity of compared methods.

Method	Runtime (ms)	Method	Params (M)	FLOPs (G)	Runtime (ms)
DCP	306.251	AOD-Net	0.002	0.214	0.850
CEP	9.182	GCANet	0.670	35.014	7.860
IDeRs	304.158	AUNet	7.417	96.924	30.230
SMIDCP	114.872	LDNet	0.029	3.672	1.160
HazeLine	1607.788	CASM	0.002	0.300	9.236
EVPM	47.704	MDNet	5.959	105.006	52.377
CAP	48.132	MABDT	11.987	121.156	212.488
		FSRNet	0.865	17.172	22.520

4.3. Interpretability Analysis of FSRNet by Feature Visualization

To validate the DFCB module's ability to distinguish hazy/clear sub-regions and clarify the interpretability of the data-driven FSRNet with atmospheric scattering principles, we conducted a feature visualization experiment. For two given hazy images, we visualized the ground-truth transmittance map (T) (darker color means denser haze) and the features of FSRNet. Specifically, we extracted feature maps from the second DFCB of each FSRM module of FSRNet in a sequence from shallow to deep layers, denoted as "F1" to "F5" in Figure 12. DFCB feature maps reflect response weights (brighter color means stronger attention to target regions). Visualization results (Figure 12) confirm two key points: First, the DFCB module effectively identifies hazy regions, indicated by the correlation between bright feature map areas and dense haze (low-transmittance) regions in T. Second, FSRNet aligns implicitly with atmospheric scattering principles. It can be observed that the shallow-layer features well capture the spatial distribution of haze. As the network deepens, features "F4" and "F5" gradually suppress haze effects and focus on the high-level semantic information of ground objects at the global scale. This demonstrates that the model implicitly learns haze-related properties to guide more effective haze removal, without relying on explicit physical priors.

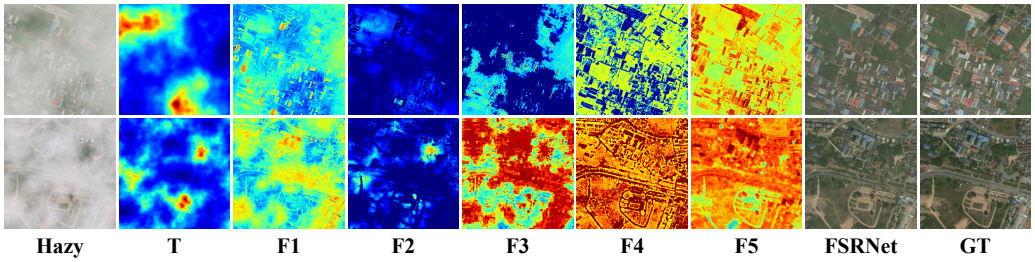


Figure 12. Feature visualization of FSRNet. From left to right: Hazy input, ground-truth transmittance map (T), DFCB feature maps (F1–F5), FSRNet dehazing result, and ground-truth (GT). T uses darker colors for denser haze, while DFCB features use brighter colors for higher response weights.

4.4. Downstream Application: Road Extraction

Haze removal is a critical preprocessing step for high-level RSI applications, so dehazing effectiveness must be verified by both visual quality and its impact on downstream pipeline tasks. Notably, visually clear images may still have subtle distortions/artifacts that degrade downstream algorithm robustness. We thus evaluate FSRNet and comparison methods on a typical RSI downstream task: road extraction. Using the RSHD dataset with masks from DeepGlobeRoad, we process hazy images with all compared dehazing algorithms and then perform road extraction by CoANet [67]. The performance is quantified by Recall, Accuracy, IoU, and F1-score.

For visual comparison, dehazing results of all methods are fused with road masks (Figure 13). In fused images, green, red, and blue pixels represent correctly extracted road regions, false positives, and false negatives, respectively. Under moderate haze (top two rows in Figure 13), DCP and IDeRs yield incomplete dehazing with obvious over-enhancement but do not severely affect the road extraction model, which still extracts most road structures. SMIDCP, however, disturbs local texture details heavily, causing severe road extraction omission errors. Under dense haze, most data-driven methods (GCANet, AUNet, MABDT, MDNet, FSRNet) achieve satisfactory dehazing results, whereas other methods fail, causing a sharp drop in downstream road extraction performance. Overall, FSRNet's reliable dehazing minimizes both false positives and false negatives in road extraction. Quantitative results in Table 11 show that while haze significantly degrades extraction accuracy, FSRNet yields the highest IoU, Recall, and F1-scores, with improvements of 66%, 78%, and 55%, respectively, over the direct extraction results on hazy images. This demonstrates that FSRNet effectively preserves essential features for high-level spatial analysis under adverse weather conditions.

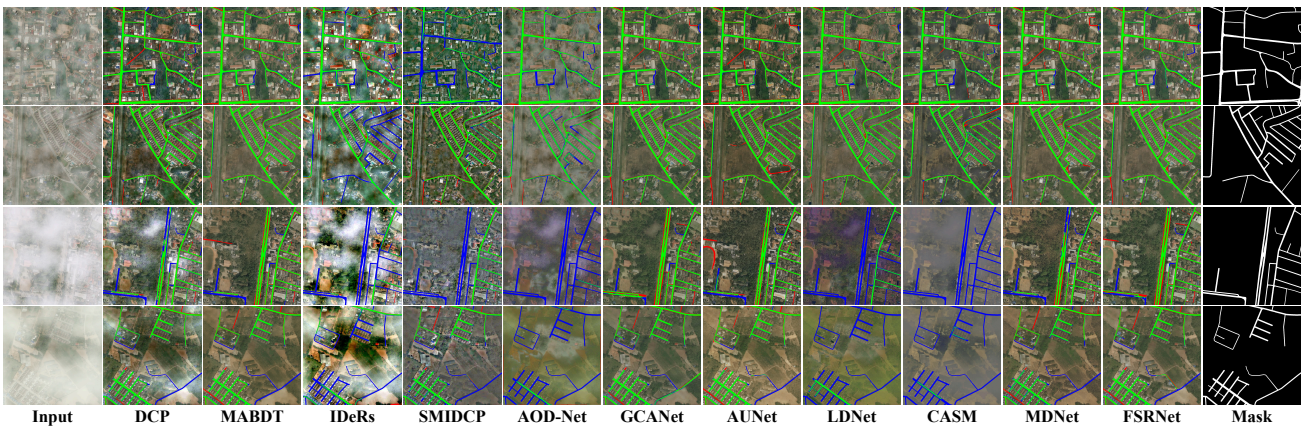


Figure 13. Visual results of the road extraction by all comparison methods on RSHD dataset. Green, red, and blue represent correctly extracted road regions, false positives, and false negatives, respectively.

Table 11. The quantitative evaluation of the road extraction results by all comparison methods on RSHD dataset.

Metrics	Hazy	DCP	MABDT	IDeRs	SMIDCP	AOD-Net	GCANet	AUNet	LDNet	CASM	MDNet	FSRNet
IoU	0.396	0.563	0.632	0.342	0.525	0.429	0.645	0.623	0.528	0.403	0.619	**0.658**
Precision	0.639	0.741	0.768	0.696	0.731	0.741	0.771	0.760	**0.789**	0.710	0.769	0.766
Recall	0.462	0.701	0.782	0.397	0.654	0.508	0.798	0.775	0.634	0.487	0.762	**0.824**
F1	0.512	0.717	0.773	0.497	0.686	0.570	0.783	0.766	0.674	0.511	0.763	**0.793**

4.5. Limitation Analysis and Future Work

Despite the promising performance of FSRNet on typical hazy remote sensing images, it has limitations in extreme scenarios, as visualized in Figure 14: (a) for sandstorm conditions, the distinct scattering characteristics of large sand particles and the severe masking of valid scene information result in ineffective restoration and pseudo-textures; (b) extremely dense haze with transmittance approaching zero ($T \approx 0$) renders the model incapable of recovering adequate underlying surface features, leading to errors in the reconstruction of the spatial structure; (c) monochromatic scenes lack sufficient chromatic constraints information, inducing over-processing artifacts and unnatural color deviations in the output; (d) the significant domain gap between remote sensing and ground-based images gives rise to dehazing failure.

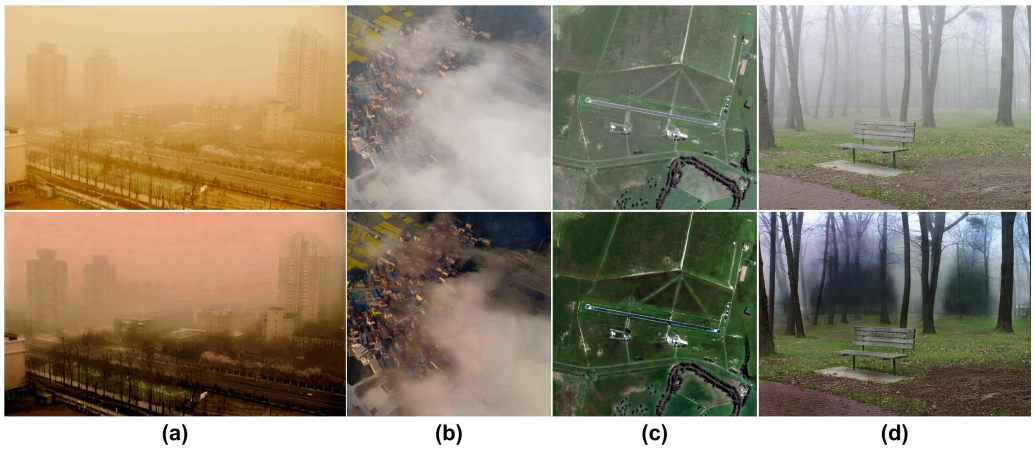


Figure 14. Performance limitations in extreme scenarios. The top row shows hazy samples, and the bottom row shows the corresponding dehazed results by the proposed FSRNet. (**a**–**d**) represent different failure scenarios.

Our future work will target these limitations to enhance the model's robustness and applicability. Specifically, we will adopt multi-modal fusion with multi-source data (e.g., infrared images, SAR data) and augment datasets with extreme scene samples to improve adaptability to harsh atmospheric conditions. Unsupervised domain adaptation techniques will be employed to bridge cross-scene domain gaps and boost generalization. Furthermore, lightweight optimization via channel pruning and knowledge distillation will facilitate practical deployment in real-world remote sensing systems, while joint training with high-level remote sensing tasks will further enhance the algorithm's practical utility for downstream applications.

5. Conclusions

In this paper, we proposed FSRNet, an innovative Feature Self-Recalibration Network tailored for remote sensing image dehazing. The network integrates a symmetric encoder–decoder with pixel shuffle/unshuffle operations for efficient multiscale feature

resampling and introduces two core components: the Dual-Stage Feature Calibration Block (DFCB), which adaptively suppresses feature redundancy through information-aware splitting and recalibration, and the Hierarchical Context-Aggregation Block (HCAB), which jointly captures global haze distribution and enhances local texture details. Extensive experiments on multiple benchmark datasets, including both synthetic and real-world remote sensing imagery, demonstrated that FSRNet consistently outperforms prior-based and deep learning-based methods across perceptual, structural, and color fidelity metrics. Particularly, FSRNet achieves superior robustness in handling non-uniform and dense haze conditions while maintaining a low computational cost.

Author Contributions: Conceptualization, Y.H. and C.L.; methodology, Y.H. and C.L.; software, Y.H. and M.S.; validation, Y.H. and X.L.; data curation, Y.H. and X.L.; writing—original draft preparation, Y.H.; writing—review and editing, M.S. and C.L.; supervision, C.L.; funding acquisition, C.L., X.L. and M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was jointly supported by the Science and Technology Key Project of the Xinjiang Production and Construction Corps (2023AB063); the Innovation Team Project of Tarim University (TDZKCX202306); the Joint Funds of Tarim University and China Agricultural University (ZNLH202505 and ZNLH202402).

Data Availability Statement: The public datasets used in the experiments are available at the following links: RRSD100 https://drive.google.com/file/d/1ntNgpte5w_bcYdL2dDUGAnHcgnSVhTXH/view?usp=drive_link (accessed on 13 November 2025), SateHaze1k https://www.dropbox.com/s/k2i3p7puuwl2g59/Haze1k.zip?dl=0 (accessed on 30 March 2025), RSHD https://drive.google.com/file/d/1ueyfp6RTY4FhsgDowyeG35yFMoHUwwac/view?usp=sharing (accessed on 30 March 2025), RSID https://drive.google.com/file/d/1FC7oSkGTthjHl2sKN-yGrKhssgV0QB4F/view?usp=sharing (accessed on 30 March 2025), and RICE https://github.com/BUPTLdy/RICE_DATASET (accessed on 30 March 2025).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [CrossRef]
2. Liu, J.; Wang, S.; Wang, X.; Ju, M.; Zhang, D. A review of remote sensing image dehazing. *Sensors* **2021**, *21*, 3926. [CrossRef]
3. Jiang, W.; Sun, Y.; Lei, L.; Kuang, G.; Ji, K. AdaptVFMs-RSCD: Advancing Remote Sensing Change Detection from binary to semantic with SAM and CLIP. *ISPRS J. Photogramm. Remote Sens.* **2025**, *230*, 304–317. [CrossRef]
4. Sun, Z.; Zhang, X.; Leng, X.; Wu, X.; Xiong, B.; Ji, K.; Kuang, G. KFIA-Net: A knowledge fusion and imbalance-aware network for multi-category SAR ship detection. *Int. J. Appl. Earth Obs. Geoinf.* **2026**, *146*, 105127. [CrossRef]
5. Zhang, X.; Zhang, S.; Sun, Z.; Liu, C.; Sun, Y.; Ji, K.; Kuang, G. Semantic scattering graph structure alignment for cross-sensor SAR image target detection. *Chin. J. Aeronaut.* **2026**, 104066. [CrossRef]
6. Wang, Z.; Liu, D.; Wang, Z.; Liao, X.; Zhang, Q. A new remote sensing change detection data augmentation method based on mosaic simulation and haze image simulation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 4579–4590. [CrossRef]
7. Narasimhan, S.G.; Nayar, S.K. Vision and the atmosphere. *Int. J. Comput. Vis.* **2002**, *48*, 233–254. [CrossRef]
8. Galdran, A.; Alvarez-Gila, A.; Bria, A.; Vazquez-Corral, J.; Bertalmío, M. On the duality between retinex and image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8212–8221.
9. Mondal, K.; Rabidas, R.; Dasgupta, R. Single image haze removal using contrast limited adaptive histogram equalization based multiscale fusion technique. *Multimed. Tools Appl.* **2024**, *83*, 15413–15438. [CrossRef]
10. Ju, M.; Ding, C.; Guo, Y.J.; Zhang, D. IDGCP: Image dehazing based on gamma correction prior. *IEEE Trans. Image Process.* **2019**, *29*, 3104–3118. [CrossRef] [PubMed]
11. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 2341–2353. [CrossRef]
12. Zhu, Q.; Mai, J.; Shao, L. A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process.* **2015**, *24*, 3522–3533. [CrossRef] [PubMed]

13. Berman, D.; Treibitz, T.; Avidan, S. Single image dehazing using haze-lines. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *42*, 720–734. [CrossRef]
14. Du, Y.; Li, J.; Sheng, Q.; Zhu, Y.; Wang, B.; Ling, X. Dehazing Network: Asymmetric Unet Based on Physical Model. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5607412. [CrossRef]
15. Wang, X.; Chen, X.; Ren, W.; Han, Z.; Fan, H.; Tang, Y.; Liu, L. Compensation atmospheric scattering model and two-branch network for single image dehazing. *IEEE Trans. Emerg. Top. Comput. Intell.* **2024**, *8*, 2880–2896. [CrossRef]
16. Lu, L.; Xiong, Q.; Xu, B.; Chu, D. Mixdehazenet: Mix structure block for image dehazing network. In Proceedings of the 2024 International Joint Conference on Neural Networks (IJCNN), Yokohama, Japan, 30 June–5 July 2024; pp. 1–10.
17. Pazhani, A.A.J.; Periyanayagi, S. A novel haze removal computing architecture for remote sensing images using multi-scale Retinex technique. *Earth Sci. Inform.* **2022**, *15*, 1147–1154. [CrossRef]
18. Khan, H.; Sharif, M.; Bibi, N.; Usman, M.; Haider, S.A.; Zainab, S.; Shah, J.H.; Bashir, Y.; Muhammad, N. Localization of radiance transformation for image dehazing in wavelet domain. *Neurocomputing* **2020**, *381*, 141–151. [CrossRef]
19. Yu, L.; Liu, X.; Liu, G. A new dehazing algorithm based on overlapped sub-block homomorphic filtering. In Proceedings of the Eighth International Conference on Machine Vision (ICMV 2015), Barcelona, Spain, 19–21 November 2015; Volume 9875, p. 987502.
20. Kumar, R.; Bhandari, A.K.; Kumar, M. Haze elimination model-based color saturation adjustment with contrast correction. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 5013610. [CrossRef]
21. Jin, R.; Wang, S.; Zhao, J.; Yang, P. Color balancing and dehazing algorithm for colored haze removal. *J. Electron. Imaging* **2020**, *29*, 043019. [CrossRef]
22. Li, J.; Hu, Q.; Ai, M. Haze and Thin Cloud Removal via Sphere Model Improved Dark Channel Prior. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 472–476. [CrossRef]
23. Ning, J.; Zhou, Y.; Liao, X.; Duo, B. Single Remote Sensing Image Dehazing Using Robust Light-Dark Prior. *Remote Sens.* **2023**, *15*, 938. [CrossRef]
24. Han, J.; Zhang, S.; Fan, N.; Ye, Z. Local patchwise minimal and maximal values prior for single optical remote sensing image dehazing. *Inf. Sci.* **2022**, *606*, 173–193. [CrossRef]
25. Liang, S.; Gao, T.; Chen, T.; Cheng, P. A remote sensing image dehazing method based on heterogeneous priors. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5619513. [CrossRef]
26. He, Y.; Li, C.; Li, X. Remote sensing image dehazing using heterogeneous atmospheric light prior. *IEEE Access* **2023**, *11*, 18805–18820. [CrossRef]
27. He, Y.; Li, C.; Bai, T. Remote sensing image haze removal based on superpixel. *Remote Sens.* **2023**, *15*, 4680. [CrossRef]
28. Bui, T.M.; Kim, W. Single image dehazing using color ellipsoid prior. *IEEE Trans. Image Process.* **2017**, *27*, 999–1009. [CrossRef] [PubMed]
29. Xu, L.; Zhao, D.; Yan, Y.; Kwong, S.; Chen, J.; Duan, L.Y. IDeRs: Iterative dehazing method for single remote sensing image. *Inf. Sci.* **2019**, *489*, 50–62. [CrossRef]
30. Li, B.; Peng, X.; Wang, Z.; Xu, J.; Feng, D. AOD-Net: All-In-One Dehazing Network. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.
31. Ullah, H.; Muhammad, K.; Irfan, M.; Anwar, S.; Sajjad, M.; Imran, A.S.; De Albuquerque, V.H.C. Light-DehazeNet: A novel lightweight CNN architecture for single image dehazing. *IEEE Trans. Image Process.* **2021**, *30*, 8968–8982. [CrossRef]
32. Zhang, Y.; Zhou, S.; Li, H. Depth information assisted collaborative mutual promotion network for single image dehazing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024; pp. 2846–2855.
33. Liu, X.; Ma, Y.; Shi, Z.; Chen, J. Griddehazenet: Attention-based multi-scale network for image dehazing. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 7314–7323.
34. Chen, D.; He, M.; Fan, Q.; Liao, J.; Zhang, L.; Hou, D.; Yuan, L.; Hua, G. Gated context aggregation network for image dehazing and deraining. In Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa Village, HI, USA, 7–11 January 2019; pp. 1375–1383.
35. Liang, S.; Gao, T.; Chen, T.; Wen, Y.; Zhang, Q.; Wang, X. INP-Net: Implicit Neural Prompting Network for Remote Sensing Image Dehazing. *IEEE Trans. Geosci. Remote Sens.* **2026**, *64*, 3649014. [CrossRef]
36. Li, Y.; Chen, X. A Coarse-to-Fine Two-Stage Attentive Network for Haze Removal of Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 1751–1755. [CrossRef]
37. Wen, Y.; Gao, T.; Li, Z.; Zhang, J.; Chen, T. Encoder-Minimal and Decoder-Minimal Framework for Remote Sensing Image Dehazing. In Proceedings of the ICASSP 2024—2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Republic of Korea, 14–19 April 2024; pp. 36–40. [CrossRef]
38. Jin, Y.; Chen, J.; Tian, F.; Hu, K. Lfd-net: Lightweight feature-interaction dehazing network for real-time remote sensing tasks. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 9139–9153. [CrossRef]

39. Li, S.; Zhou, Y.; Ren, W.; Xiang, W. PFONet: A progressive feedback optimization network for lightweight single image dehazing. *IEEE Trans. Image Process.* **2023**, *32*, 6558–6569. [CrossRef] [PubMed]
40. Li, S.; Zhou, Y.; Kung, S.Y. PSRNet: A Progressive Self-Refine Network for Lightweight Optical Remote Sensing Image Dehazing. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5648413. [CrossRef]
41. Zhang, Y.; Luo, X.; Sun, Q.; Wang, Y.; Qu, H.; Yi, Z. Retina-Inspired Lightweight Spiking Convolutional Neural Network for Single-Image Dehazing. *IEEE Trans. Neural Netw. Learn. Syst.* **2025**, *36*, 12580–12594. [CrossRef]
42. Pushpalatha, D.; Prithvi, P. LDFormer: Lightweight dehazing transformer. *Neural Comput. Appl.* **2025**, *37*, 20049–20068. [CrossRef]
43. He, Y.; Li, C.; Li, X.; Bai, T. A Lightweight CNN Based on Axial Depthwise Convolution and Hybrid Attention for Remote Sensing Image Dehazing. *Remote Sens.* **2024**, *16*, 2822. [CrossRef]
44. Zhong, S.; Zhao, X.; Liu, D.; Su, H.; Xie, Z.; Fan, B. High-resolution, lightweight remote sensing via harmonic diffractive optical imaging systems and deep denoiser prior image restoration. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5621717. [CrossRef]
45. Nguyen, B.A.; Kha, M.B.; Dao, D.M.; Nguyen, H.K.; Nguyen, M.D.; Nguyen, T.V.; Rathnayake, N.; Hoshino, Y.; Dang, T.L. UFR-GAN: A lightweight multi-degradation image restoration model. *Pattern Recognit. Lett.* **2025**, *197*, 282–287. [CrossRef]
46. Ye, T.; Zhang, Y.; Jiang, M.; Chen, L.; Liu, Y.; Chen, S.; Chen, E. Perceiving and modeling density for image dehazing. In *Proceedings of the European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 130–145.
47. Song, W.; Gao, Y.; Xiong, J.; Lin, H.; Li, D.; Zhang, Y. Two Stage Dehazing Framework for Dense and Non-Homogeneous Dehazing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024; pp. 6722–6731.
48. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
49. Shi, W.; Caballero, J.; Huszár, F.; Totz, J.; Aitken, A.P.; Bishop, R.; Rueckert, D.; Wang, Z. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1874–1883.
50. Zhang, Q.; Jiang, Z.; Lu, Q.; Zeng, Z.; Gao, S.H.; Men, A. Split to be slim: An overlooked redundancy in vanilla convolution. In Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence, Yokohama, Japan, 7–15 January 2021; pp. 3195–3201.
51. Li, J.; Wen, Y.; He, L. Scconv: Spatial and channel reconstruction convolution for feature redundancy. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 6153–6162.
52. Hu, J.; Shen, L.; Sun, G. 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; pp. 7132–7141.
53. 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; pp. 286–301.
54. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
55. Ning, J.; Yin, J.; Deng, F.; Xie, L. MABDT: Multi-scale attention boosted deformable transformer for remote sensing image dehazing. *Signal Process.* **2025**, *229*, 109768. [CrossRef]
56. Huang, B.; Zhi, L.; Yang, C.; Sun, F.; Song, Y. Single satellite optical imagery dehazing using SAR image prior based on conditional generative adversarial networks. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Snowmass Village, CO, USA, 1–5 March 2020; pp. 1806–1813.
57. Chi, K.; Yuan, Y.; Wang, Q. Trinity-Net: Gradient-guided Swin transformer-based remote sensing image dehazing and beyond. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4702914. [CrossRef]
58. Lin, D.; Xu, G.; Wang, X.; Wang, Y.; Sun, X.; Fu, K. A remote sensing image dataset for cloud removal. *arXiv* **2019**, arXiv:1901.00600. [CrossRef]
59. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 586–595.
60. Zhang, L.; Zhang, L.; Mou, X.; Zhang, D. FSIM: A feature similarity index for image quality assessment. *IEEE Trans. Image Process.* **2011**, *20*, 2378–2386. [CrossRef] [PubMed]
61. Xia, G.S.; Hu, J.; Hu, F.; Shi, B.; Bai, X.; Zhong, Y.; Zhang, L.; Lu, X. AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3965–3981. [CrossRef]
62. Li, K.; Wan, G.; Cheng, G.; Meng, L.; Han, J. Object detection in optical remote sensing images: A survey and a new benchmark. *ISPRS J. Photogramm. Remote Sens.* **2020**, *159*, 296–307. [CrossRef]
63. Pan, X.; Xie, F.; Jiang, Z.; Shi, Z.; Luo, X. No-reference assessment on haze for remote-sensing images. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1855–1859. [CrossRef]

64. Min, X.; Zhai, G.; Gu, K.; Yang, X.; Guan, X. Objective quality evaluation of dehazed images. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 2879–2892. [CrossRef]
65. Zhou, W.; Zhang, R.; Li, L.; Yue, G.; Gong, J.; Chen, H.; Liu, H. Dehazed image quality evaluation: From partial discrepancy to blind perception. *IEEE Trans. Intell. Veh.* **2024**, *9*, 3843–3858. [CrossRef]
66. Guan, T.; Li, C.; Gu, K.; Liu, H.; Zheng, Y.; Wu, X.j. Visibility and distortion measurement for no-reference dehazed image quality assessment via complex contourlet transform. *IEEE Trans. Multimed.* **2022**, *25*, 3934–3949. [CrossRef]
67. Mei, J.; Li, R.J.; Gao, W.; Cheng, M.M. CoANet: Connectivity attention network for road extraction from satellite imagery. *IEEE Trans. Image Process.* **2021**, *30*, 8540–8552. [CrossRef] [PubMed]

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.