Article

Towards a Novel Generative Adversarial Network-Based Framework for Remote Sensing Image Demosaicking

Yuxuan Guo [1,2] , Xuemin Zhang [1,2,3,*] and Guang Jin [1,2,3]

1   School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China;
    guoyuxuan@whu.edu.cn (Y.G.); ht004429@whu.edu.cn (G.J.)
2   Institute of Aerospace Science and Technology, Wuhan University, Wuhan 430072, China
3   Hubei Luojia Laboratory, Wuhan 430070, China
*   Correspondence: zhangxuemin@whu.edu.cn

Abstract: During satellite remote sensing imaging, the use of Bayer mode sensors holds significant importance in saving airborne computing resources and reducing the burden of satellite transmission systems. The demosaicing techniques play a key role in this process. The integration of Generative Adversarial Networks (GANs) has garnered significant interest in the realm of image demosaicking, owing to their ability to generate intricate details. However, when demosaicing mosaic images in remote sensing techniques, GANs, although capable of generating rich details, often introduce unpleasant artifacts while generating content. To address this challenge and differentiate between undesirable artifacts and realistic details, we have devised a novel framework based on a Progressive Discrimination Strategy within a Generative Adversarial Network architecture for image demosaicking. Our approach incorporates an artifact-weighted Location Map refinement technique to guide the optimization process towards generating authentic details in a stable and precise manner. Furthermore, our framework integrates a global attention mechanism to boost the interaction of spatial-channel information across different dimensions, thereby enhancing the performance of the generator network. Moreover, we conduct a comparative analysis of various prevalent attention mechanisms in the context of remote sensing image demosaicking. The experimental findings unequivocally demonstrate that our proposed methodology not only achieves superior reconstruction accuracy on the dataset but also enhances the perceptual quality of the generated images. By effectively mitigating artifacts and emphasizing the generation of true details, our approach represents a significant advancement in the field of remote sensing image demosaicking, promising enhanced visual fidelity and realism in reconstructed images.

Keywords: remote sensing image demosaicking; generative adversarial networks; attention mechanism

1. Introduction

The rapid evolution of digital imaging technology has significantly enhanced the data rate and resolution of remote-sensing satellite images. This surge in acquired data has underscored the critical importance of onboard storage capacity and transmission bandwidth. Capturing comprehensive satellite images involves the simultaneous acquisition of red (R), green (G), and blue (B) information for each pixel, necessitating the use of sensors equipped with distinct color filters. However, this approach not only escalates hardware costs and dimensions but also engulfs storage capacity and transmission bandwidth with sensor data. Consequently, an increasing number of satellites, such as Jilin-1 and Zhuhai-1, are opting for imaging systems that employ single CCD and CMOS sensors with a Color Filter Array (CFA) [1], effectively reducing the data volume per image by two-thirds and alleviating the strain on satellite-to-ground transmission. The Bayer pattern emerges as the predominant CFA arrangement, strategically interleaving red, green, and blue color filters in a pixel-wise alternating manner, with a higher prevalence of green filters compared to red (or blue) filters. This configuration aligns well with human color perception and is often referred to as

a mosaic template due to its design. Leveraging the Bayer pattern enables satellite camera systems to maintain image quality while significantly increasing the quantity of efficiently acquired data. This optimization is crucial for maximizing onboard resource utilization and mitigating the pressure on satellite-to-ground transmission systems. The interpolation of satellite images in Bayer mode serves as a crucial initial step in the processing workflow, directly impacting the overall quality of satellite imagery. Given that each pixel in a Bayer image provides only one color measurement value, the reconstruction of a full-color image through Bayer interpolation during ground processing, also known as CFA interpolation or demosaicking [2], becomes imperative.

Numerous image demosaicking methods have been proposed in the past decades. Traditional model-based methods for image demosaicking include interpolation-based methods [3–16], dictionary-based methods [17,18], and iterative methods [19]. Traditional interpolation methods initially performed color image reconstruction by calculating gray gradients and color differences. Later, residual interpolation (RI)–based demosaicking methods were shown to be superior to color component difference-based interpolation methods. Iterative methods [19] solved the problem that existing RI-based methods could only adequately recover the R and B channels, and high precision G channels could be reconstructed by an iterative RI (IRI) process, but such methods increase the computational effort significantly. The common drawback of the above traditional methods is that the model parameters are usually determined by hand and lack generalization to optimize color data sets with different characteristics.

Data-driven convolutional neural network-based (CNN-based) approaches have significant advantages over traditional model-based approaches [20–29]. The goal of the data-driven approach is to learn a nonlinear mapping from the input image space to the output image space rather than analyzing the global optimization problem to obtain an optimal solution, thus providing a novel idea of end-to-end optimization. A CNN-based denoising convolutional neural network (DnCNN) [30] was proposed to independently perform demosaicking and denoising, which effectively suppresses the appearance of noise. Due to the potential of generating rich details, generative adversarial networks (GANs) have been gradually introduced into the field of image demosaicking in recent years. Joint demosaicking and denoising networks based on generative adversarial networks (GANs) enhance the demosaicking effect by using a combination of perceptual and adversarial loss functions. However, due to the instability of GAN network training, the demosaicking process more or less introduces many visual artifacts that do not match the real scene.

As shown in Figure 1, the demosaicking process has different degrees of true detail recovery for different areas of the image. For smooth areas, the resulting rich detail can restore the true information of the image to a large extent. However, for many fine regular structures or sharp transitions between adjacent pixels, demosaicking often produces unpleasant artifacts. Existing demosaicking networks do not distinguish between real details and artifacts but train both in the same general way of learning, and also lack consideration of information about the spatial-channel attention interaction dimension, and thus are often prone to distortions such as color artifacts and zippers to varying degrees during training. Moreover, in contrast to natural images, remote sensing images are typically captured from a distant, aerial perspective, resulting in a prevalence of low-frequency regions. Consequently, the applicability of current algorithms designed for demosaicing natural images to the domain of remote sensing images warrants further investigation and verification.

To address the above problem, we propose a novel end-to-end discriminant-enhanced generation framework that uses a global attention-based generative adversarial network for remote sensing image demosaicking, along with progressive discriminative learning to distinguish unpleasant visual artifacts from real details to normalize adversarial training. Our work has the following three main contributions. In order to obtain high-quality images with rich fidelity details, we develop a new discriminant-enhanced generative adversarial networks (GANs)-based image demosaicking method to achieve end-to-end

full-color image reconstruction. To distinguish visual artifacts from true details, we employ a progressive discrimination framework that explicitly penalizes artifacts without sacrificing true details by refine artifact weighting-location maps, while optimizing the perceptual function to improve the generation details. To enhance the spatial-channel information interaction across latitudes, we introduce a global attention mechanism in the generator to improve the network performance.
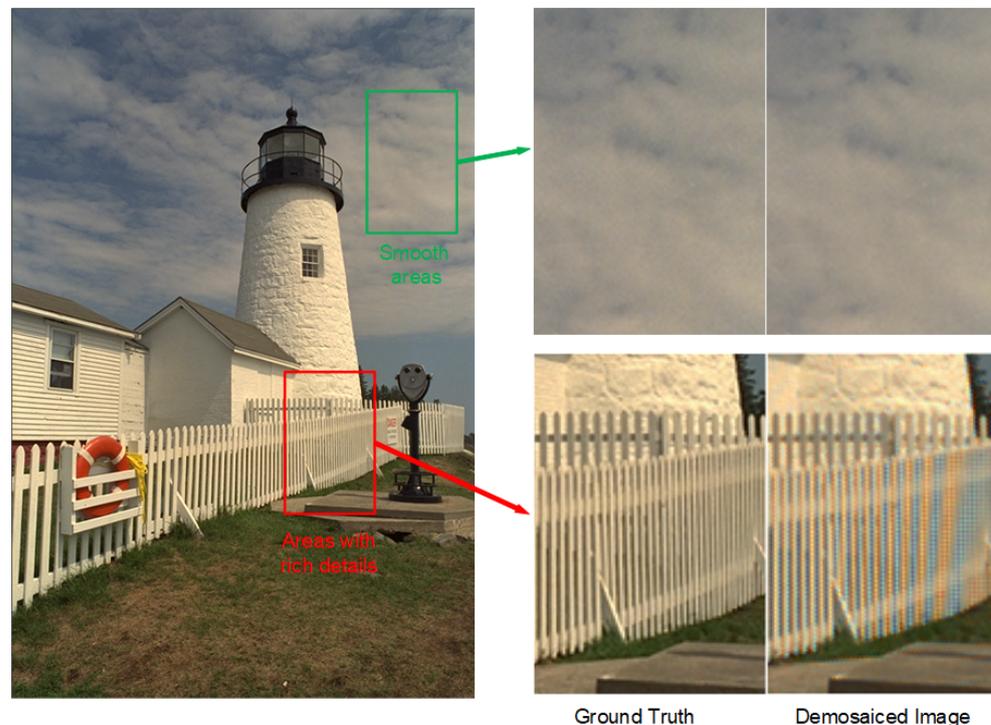

Figure 1. The demosaicking process has different levels of true detail reproduction for different areas of the image (smooth areas are well restored, while areas with rich and dense detail are poorly restored).

We compared the performance of several currently dominant attention mechanisms in the demosaicking task to identify the most suitable attention mechanism for the task. In parallel, we conducted extensive experiments on the dataset. Quantitative and qualitative comparisons show that our approach performs significantly better than the state-of-the-art works.

2. Related Work

Image demosaicking is an essential step in the image signal processing pipeline that has been well studied. Therefore, numerous methods have been proposed in the past decades. The methods for image demosaicking generally include model-based algorithms and data-driven algorithms. In recent years, some studies have also transitioned from color demosaicking of natural images to demosaicking specifically for remote sensing images. In this section, we review image demosaicking and discuss the different proposed solutions separately.

2.1. Model-Based Methods

Intuitively, image demosaicking can be best understood as an extension of image interpolation for grayscale images. Due to the high correlation between the three color channels, Sakamoto et al. [31] found that a linear interpolation method that considers the correlation of G determined using R/B pixels can restore colors well and reduce the appearance of artifacts. Hua et al. [32] proposed an interpolation algorithm that first estimates the green component, then calculates the chromatic aberration images (R-G and B-G) and uses compensation for the interpolated G and edge adaptation method to interpolate

the full-resolution chromatic aberration images. Chung et al. [33] proposed a method that combines traditional chromaticity subsampling with distortion minimization-based Luma modification to improve the quality of the reconstructed RGB full-color images. Meanwhile, to make full use of spatial information to reduce edge blurring, Sher et al. [34] introduced a new image interpolation using the spatial relationship between adjacent pixels. Menon et al. [35] proposed a new demosaicking technique based on directional filtering and a posteriori decision-making. Nallaperumal et al. [36] used a novel adaptive edge protection, edge-directed interpolation technique to reproduce the color of Bayer mosaic images. Yang et al. [37] proposed a block edge estimation method considering all color channels. Sun et al. [38] proposed a hybrid demosaicking algorithm based on fuzzy edge strength and residual interpolation (RI) methods. In terms of interpolation direction, Fan et al. [39] demonstrated that an omnidirectional estimation-based demosaicking algorithm outperforms horizontal and vertical estimation methods. In addition, considering the algorithmic efficiency and complexity, Su et al. [40] proposed an efficient color interpolation using simple pixel-level data-dependent triangulation. Karch et al. [41] developed a fast SR method based on the adaptive wiener filter (AWF) super-resolution (SR) algorithm and using a global channel-to-channel statistical model. Addressing common issues such as zipper artifacts, noise, and false colors at edges that often arise during Bayer interpolation reconstruction of video satellite images, Wu et al. [42] introduced an enhanced filtering method, which is founded on the reconstruction of luminance and chrominance signals, offering a valuable approach for the further processing and utilization of satellite video data. Although traditional methods are able to restore the true information to some extent, the drawback is that these methods are based on the manual derivation of the optimal parameters from the model, are not generalizable, and are not guaranteed to be applicable to other new data.

2.2. Data-Driven Methods

Even complex demosaicking tasks can be replaced by a single end-to-end deep learning model without prior knowledge of the interpolation algorithms and parameters used for the processing. Syu et al. [24] proposed two CNN models to learn end-to-end mapping relationships between mosaic samples and original image patches with complete information. Zhang et al. [43] used deep convolutional neural networks to learn CFA and demosaicking, optimizing the spectral sensitivity function (SSF) while considering filter alignment. Tang et al. [44] formulated demosaicking as a recovery problem and solved it by minimizing the difference between the input original image and the sampled panchromatic result using a CNN-based approach. Ignatov et al. [45] proposed PyNET based on a pyramidal CNN architecture that capable of performing all ISP steps, such as image demosaicking, denoising, white balance, color and contrast correction, to convert RAW Bayer data obtained directly from a mobile camera sensor to photos taken with a professional high-end DSLR camera. Sharif et al. [46] proposed a new spatially asymmetric attention module to jointly learn bidirectional transformations and large kernel global attention to reduce visual artifacts, addressing the challenge of learning RGB image reconstruction from noisy Nona-Bayer CFA. Later, generative adversarial networks (GANs) were also introduced to the field of image demosaicking due to the advantage of generating rich details. Weisheng Dong et al. [47] proposed a joint demosaicking and denoising network based on generative adversarial networks (GANs) and used a combination of perceptual and adversarial loss functions to enhance the demosaicking effect. Meanwhile, there is an increasing interest in joint demosaicking and denoising tasks. Huang et al. [48] proposed a lightweight convolutional neural network for the joint demosaicking and denoising (JDD) problem. Park et al. [49] proposed a variational deep image prior network for joint demosaicking and denoising that can be trained on a single patterned image for patterned images with varying degrees of noise. Khadidos et al. [50] first denoised the mosaic images using a CNN and then demosaicked them using the residual learning strategy of a single specialized network. However, the demosaicking process tends to introduce many unpleas-

ant visual artifacts due to the instability of network training and the lack of consideration of spatial-channel attention interaction dimension information.

To summarize, model-based methods mainly rely on manually deriving the optimal parameters of the model, relying on a priori knowledge and specific algorithms; whereas data-driven methods utilize deep learning models without much a priori knowledge and learn mapping relationships through data. Model-based approaches can recover real information to a certain extent. However, they need to manually tune the parameters, are less versatile and their applicability to new data is not guaranteed, and the algorithm efficiency and complexity may be limited. Data-driven methods can handle complex tasks, are adaptable through end-to-end training, and can generate rich details. However, unstable network training may introduce more visual imperfections and may lack sufficient consideration of information on spatial-channel attention interaction dimensions.

3. Method

Since the demosaicking process has different fidelity in restoring different areas of the image. Smooth areas with less detailed textures can be well restored with true details, while areas rich in high-frequency information often introduce unpleasant visual artifacts. Therefore, the two cannot be simply treated uniformly. To be able to distinguish artifacts from real details, this paper proposes a new progressive discriminative strategy to regulate the optimization of the network toward a more stable and accurate direction by constructing an artifact weighting map [51]. Figure 2 illustrates the generative adversarial network framework with a progressive discriminative strategy, where we add a discriminant-enhanced learning link in addition to the initial GAN.
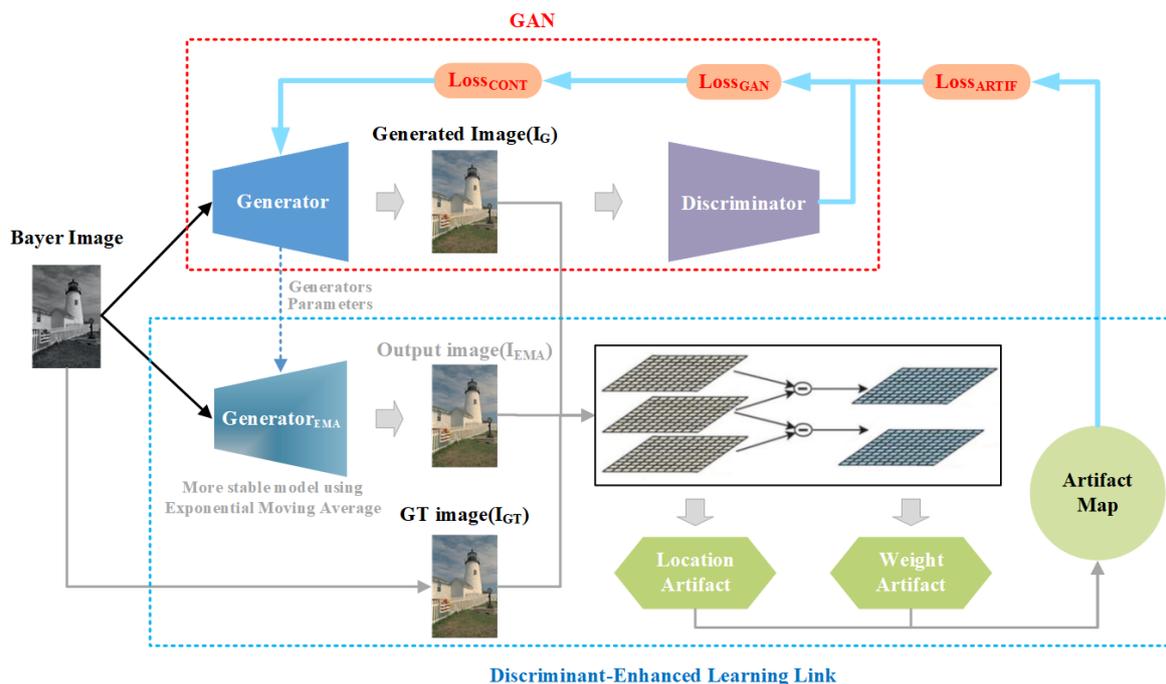

Figure 2. Generative adversarial network framework with a progressive discriminative strategy.

To address this problem, we use the exponential sliding average technique to obtain a network that generates smoother images, because the exponential sliding average is the average of the values taken by the variables over a period of history, which is less oscillating than the original network and does not make the sliding average fluctuate greatly due to an anomaly in one of the values taken. Using ϕ to denote the generator model optimized

by dynamic gradient descent, we use the exponential moving average (EMA) technique to temporarily integrate the more stable model ϕ_{EMA} from ϕ as follows:

$$\phi_{EMA}^{(k)} = \alpha \cdot \phi_{EMA}^{(k-1)} + (1 - \alpha)\phi^{(k)} \tag{1}$$

where α in Equation (1) represent the weighting parameters. Pixels are considered reasonable only if the residual between the generated image, and the true image is less than the residual between the exponential sliding average network output image and the true image. Others are considered as outliers that appear as unpleasant visual artifacts. This move yields the location of the artifacts in the weighted map.

Since the variance can reflect the stability of the data over a period of time, we use the local variance of the residuals between the generated and true images to describe the difference between artifacts and true details and further refine it to obtain a finer weight map. As shown in Figure 3, the pixel-level weight map is characterized by the local variance of the residual map in a 7×7 neighborhood for the artifactual pixels:

$$M(i, j) = \underset{7 \times 7}{V}(I_{GT} - I_G), \tag{2}$$

where V represents the variance, pixel-level weight maps can effectively detect artifactual pixels in smooth regions. However, since the local variance is computed within a very small perceptual field, distinguishing artifacts from edges and textures is unstable and can be overly responsive to pixels in some detail-rich regions. Therefore, we need further refinement to obtain a stable patch-level weight map $\sigma M(i, j)$, and compute a stable patch-level variance σ based on the whole residual map as follows:

$$\delta = (V(I_{GT} - I_G))^{1/a} \tag{3}$$

The final weight map $M_{artifact}$ is updated for the anomalous pixels of $|I_{GT} - I_G| \geq |I_{GT} - I_{EMA}|$:

$$M_{artifact}(i, j) = \begin{cases} 0, & if |I_{GT} - I_G| < \\ & |I_{GT} - I_{EMA}| \\ \delta \cdot M(i, j), & if |I_{GT} - I_G| \geq \\ & |I_{GT} - I_{EMA}| \end{cases} \tag{4}$$
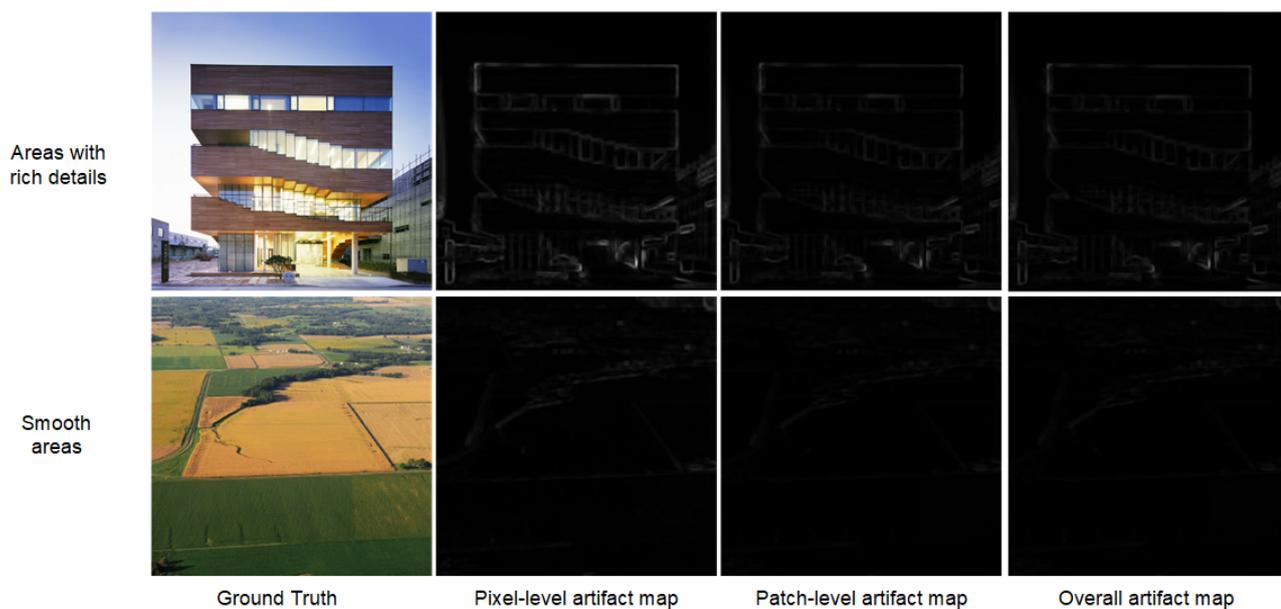

Figure 3. Visualization of the artifact map.

In summary, this paper employs a new progressive discriminative strategy to distinguish unpleasant visual artifacts from real details, and our network can better learn the features of real details through discriminative augmentation learning. Unlike $loss_{GAN}$, our discriminative enhancement strategy can better tune the model training and can consistently generate perceptually realistic details while suppressing visual artifacts.

3.1. Generator Network

The purpose of this generator is to convert a single-channel Bayer image into a full-color image with 3-channel output. It contains 1 compressed extraction block, 16 GAB blocks, 1 upsampling block and 2 convolutional layers. To address the problem that the demosaicking task does not consider spectral–spatial cross-dimensional information, we use a generator based on multiple GAB blocks to fuse local and global attention information while preserving long-hop and short-hop connections to facilitate the flow of information inside and outside the module. Experiments demonstrate that a broader network designed based on multiple GAB blocks does help to improve our demosaicking performance. To implement the above network, we propose an attention block using short hops and a global attention mechanism (as in Figure 4). Specifically, we use two convolutional layers with 64 feature maps and a batch normalization layer, and then use ParametricReLU as the activation function [52]. Then, we use the global attention mechanism (GAM) [53] to amplify the cross-dimensional spatial-channel dependence and enhance spatial information fusion. The GAB module proposed in this paper can recover two-thirds of the lost data in the spectrum, as well as the lost high-frequency details.
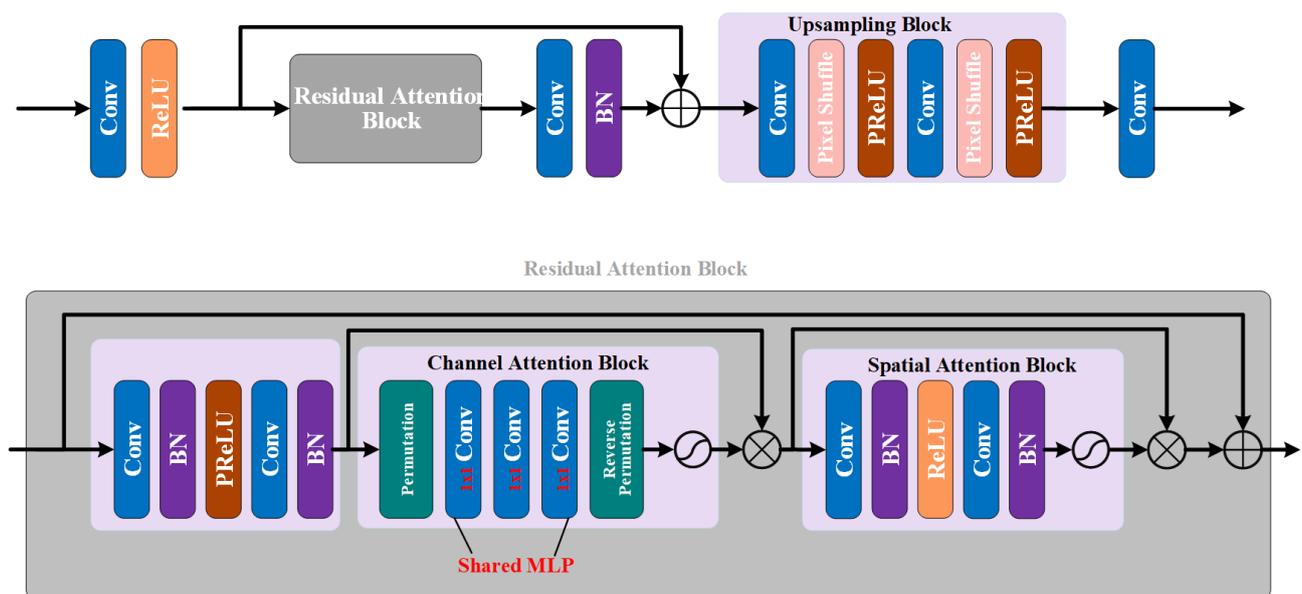

Figure 4. Structure of the generator network.

3.2. Discriminator Network

We also use a discriminator to distinguish real color images from those synthesized by the generator, as shown in Figure 5, which contains 8 convolutional layers, using LeakyReLU activation, and the number of 3×3 filter kernels is increasing from 2^6 to 2^9 kernels. The number of features is increased while using layered convolution to reduce the image resolution. Finally, two dense layers and a sigmoid activation function are used to obtain the discriminative results [52].
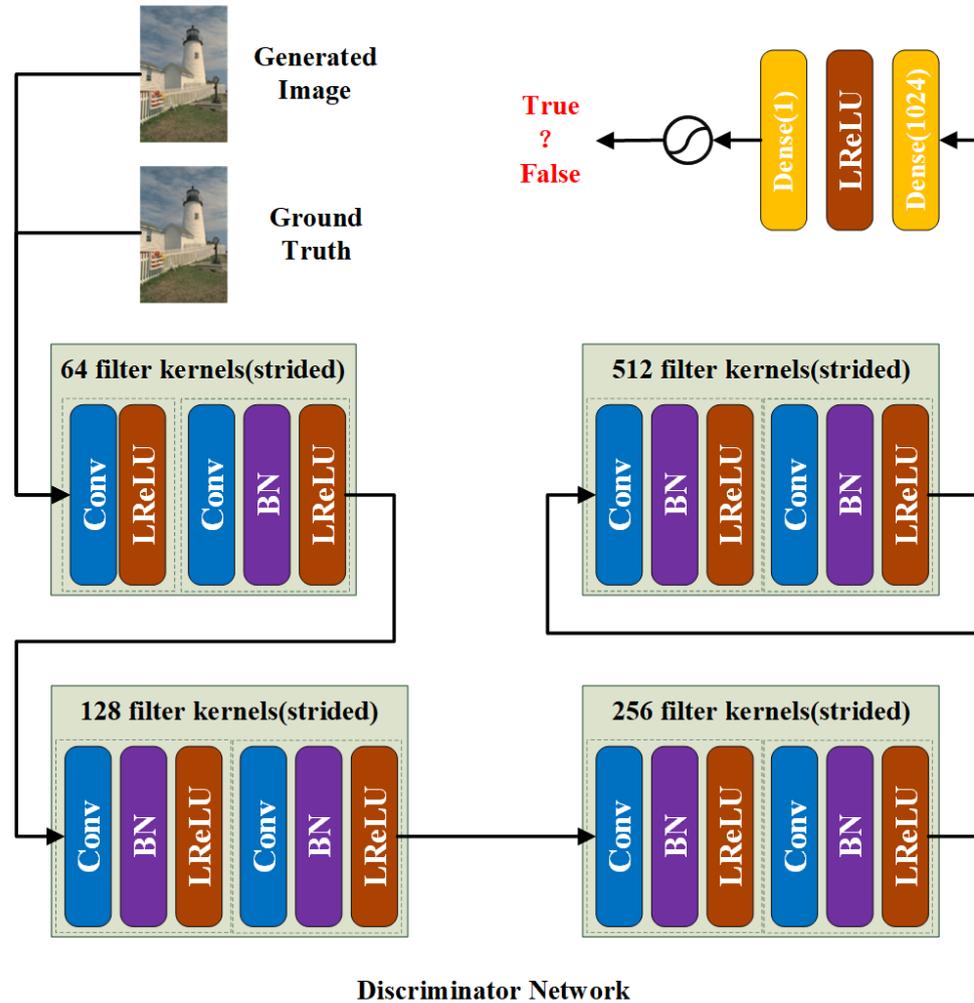
Figure 5. Structure of the discriminator network.

3.3. Loss Function

The purpose of demosaicking is to recover the true color information and high-frequency details of the image and to reduce the artifacts. Therefore, we describe the loss function of the mosaic network in terms of the weighted sum of content loss and adversarial loss, i.e.,

$$L_{LDL}^{DEMSC} = L_{MSE}^{DEMSC} + L_{VGG}^{DEMSC} + L_{ARTIF}^{DEMSC} + 10^{-3}L_{GAN}^{DEMSC} \tag{5}$$

One of the content losses consists of pixel loss, feature loss and artifact discrimination loss. The pixel-level loss is optimized by calculating the MSE. The feature loss uses the ReLU activation layer [54] of a pre-trained 19-layer VGG network to define a loss function close to the perceptual similarity.

To better learn the features of real details to suppress visual artifacts, we define the artifact discriminative loss [47] as follows:

$$L_{ARTIF}^{DEMSC} = \parallel M_{artifact} \cdot (I_{GT} - I_G) \parallel_1 \tag{6}$$

Artifact discrimination loss optimizes our network toward learning real details to suppress visual artifacts.

4. Results

4.1. Datasets and Implementation Details

One of the primary challenges encountered in data-driven demosaicking methods lies in the scarcity of real datasets containing mosaic original images paired with ground truth RGB images. Consequently, many existing approaches resort to training demosaicking networks using artificially generated data stitched from existing RGB images. In our study, we trained our proposed network using the VOC2012 [55] dataset, comprising 16,700 training images and 426 validation images. To augment the original training data, we expanded the dataset to include 200,000 training images, all of which were randomly cropped into patches measuring 88×88 pixels.

For the test phase, we employed three natural image datasets—Kodak, McMaster [56], and Set5—as well as two remotely sensed image datasets, namely SateHaze1k and DOTA-v1.0. The Kodak dataset comprises 24 images, each with a resolution of 768×512, while the McMaster dataset consists of 18 images extracted from 500×500 high-resolution images. The Set5 dataset includes 5 images. DOTA-v1.0 [57], the largest optical remote sensing image dataset to date, is a collaborative effort involving the State Laboratory of Remote Sensing at WU and the School of Telecommunication at HUST, among others. This dataset, sourced from Google Earth, as well as the Chinese satellites GF-2 and JL-1, encompasses 2806 remote sensing images with sizes ranging from 800×800 to 4000×4000 pixels. It comprises a total of 188,282 instances categorized into 15 classes, such as airplanes, boats, baseball infields, and various other objects. The SateHaze1k dataset [58], primarily utilized for image generation and denoising tasks, originates from the GF-2 and GF-3 satellites and was released by Tsinghua University in 2017. The visible images in this dataset are sized at $512 \times 512 \times 3$. To prepare the training and test data, we applied Bicubic interpolation to downsample the original high-resolution images by a factor of 2 before generating the "RGGB" Bayer pattern.

In all experimental setups, our models were developed using the PyTorch framework and executed on an NVIDIA GeForce RTX 2080Ti GPU. The batch size for training patches was fixed at 32, and our novel network architecture underwent training for a total of 500 epochs. We opted for the Adam optimizer with $\beta_1 = 0.9$ and $\beta_2 = 0.999$ parameters. The initial learning rate was initialized to 0.001, with a default decay factor of 0. Within our proposed GAN network, we configured the GAB blocks to a count of 16. The majority of convolutional layers employed a kernel size of 3×3, each with 64 filters. The reduction ratio was established at $r = 4$. The final layer was equipped with 3 filters to produce complete color image outputs.

4.2. Experiment Results

In this section, we present a comparative analysis of the proposed demosaicking method against several existing techniques, including Bicubic interpolation, FlexISP [59], ADMM [60], JDSR [61], and DPIR [62,63]. The evaluation is conducted on the Kodak, McMaster, Set5, SateHaze1k, and DOTA-v1.0 datasets, with Table 1 displaying the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) metrics for each method. PSNR is a widely adopted objective metric for image evaluation, quantifying the error between corresponding pixel values, with values exceeding 40 dB indicative of excellent image quality. However, as PSNR may not align perfectly with the characteristics of the human visual system (HVS) [64], the Structural Similarity Index (SSIM) was introduced to capture structural attributes like brightness, contrast, and image structure. SSIM leverages the strong correlation among neighboring pixels to assess the structural coherence of objects within an image, making it a valuable tool for evaluating structural fidelity. While PSNR reflects the overall image quality, SSIM is particularly adept at discerning distortions in demosaicked images. Our results reveal that the proposed method outperforms all others in both PSNR and SSIM metrics, underscoring its ability to produce visually consistent results that faithfully capture the perceptual structure of the reconstructed images.

Table 1. PSNR (peak signal-to-noise ratio) (dB) and SSIM (structural similarity) results. The comparison methods are evaluated on the Kodak, McMaster, Set5, SateHaze1k, and DOTA-v1.0 datasets.

Method	McMaster		Kodak		Set5		SateHaze1k		DOTA-v1.0	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bicubic	34.331	0.9790	34.564	0.9799	36.139	0.9827	26.293	0.9336	36.658	0.9826
FlexISP	34.938	0.9767	35.113	0.9709	37.215	0.9771	23.435	0.8265	40.079	0.9831
ADMM	32.370	0.9575	31.481	0.9382	32.566	0.9638	19.646	0.6176	32.798	0.9169
JDSR	35.968	0.9845	40.677	0.9913	36.215	0.9839	37.419	0.9921	36.837	0.9861
DPIR	37.832	0.9885	40.650	0.9915	39.526	0.9837	35.803	0.9907	43.742	0.9893
RSDM-GAN (OURS)	**39.394**	**0.9914**	**41.411**	**0.9937**	**39.581**	**0.9865**	**37.232**	**0.9943**	**45.446**	**0.9908**

Figures 6 and 7 present both qualitative and quantitative comparisons of various algorithms on natural image datasets. Figure 6 illustrates a line graph showcasing the PSNR and SSIM values of different demosaicking algorithms on the Kodak and McMaster datasets. The results in Figure 6 indicate that the proposed algorithm outperforms on the majority of images in terms of both PSNR and SSIM metrics.

(**a**) PSNR of Kodak

(**b**) SSIM of Kodak

(**c**) PSNR of McMaster
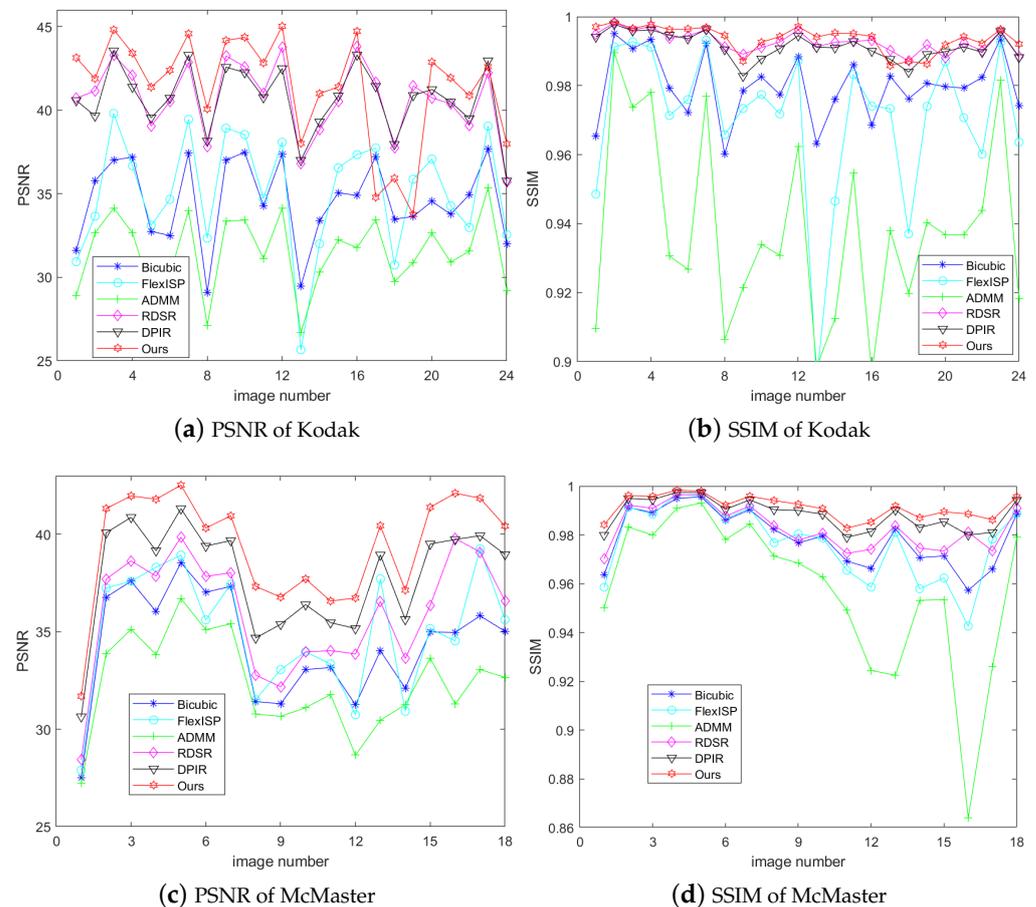
(**d**) SSIM of McMaster

Figure 6. Performance evaluation of demosaicking methods on Kodak and McMaster Datasets: PSNR and SSIM comparison

To quantitatively compare the demosaicking performance of the proposed method against state-of-the-art algorithms, we randomly selected two scenarios from the Set5 dataset and displayed them in Figure 7. In the first scene, different algorithms exhibit varying degrees of smoothing on the spots present on the child's cheek, with JDSR, FlexISP, and ADMM showing excessive smoothing. A careful comparison reveals that our algorithm preserves finer details of the facial skin spots more faithfully to the Ground Truth. In the second scene, the Bicubic algorithm appears to introduce more color aliasing compared

to other methods, while the ADMM algorithm also exhibits color aliasing, a more critical issue is the significant loss of details due to excessive smoothing. JDSR and FlexISP suffer from similar drawbacks. Although the DPIR algorithm does not exhibit the aforementioned issues, it falls short in color restoration compared to our proposed method, a distinction that becomes apparent upon closer inspection. In conclusion, through qualitative and quantitative comparisons of different demosaick methods on natural image datasets, it is evident that the results of our proposed method are more realistic, natural, and closer to ground truth than those of other algorithms. Consequently, our method outperforms existing techniques in terms of performance and fidelity.
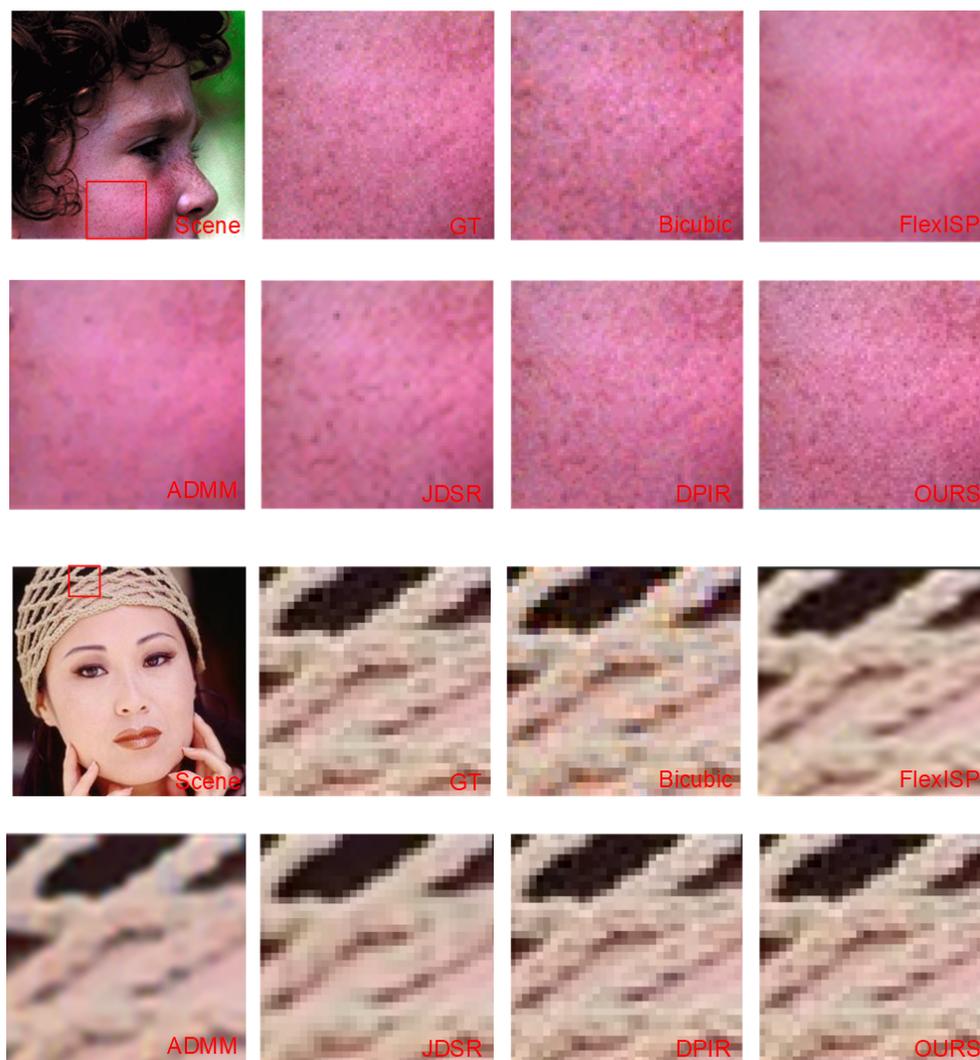

Figure 7. Comparison of the effects of different demosaicking methods for the Set5 dataset.

To qualitatively and quantitatively evaluate the performance of different demosaicking algorithms on remote sensing datasets, we, respectively, randomly selected two scenarios from the SateHaze1k and DOTA-v1.0 datasets.

The qualitative assessment results in terms of PSNR and SSIM are presented in Table 2, with corresponding subjective comparison images displayed in Figures 8 and 9. The data in Table 2 demonstrates that the proposed method achieves the best metrics on all randomly selected images. In the evaluation process, the selected scenarios from the remote sensing datasets were carefully analyzed to assess the effectiveness of each demosaicking algorithm. The results revealed that our proposed method consistently outperformed the competing algorithms in terms of both objective metrics (PSNR and SSIM) and subjective visual quality.

Table 2. PSNR (peak signal-to-noise ratio) (dB) and SSIM (structural similarity) results. The comparison methods are evaluated on the SateHaze1k and DOTA-v1.0 datasets.

Image	Bicubic		FlexISP		ADMM		JDSR		DPIR		OURS	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
1	28.004	0.9564	27.617	0.9320	25.968	0.8904	37.374	0.9940	36.245	0.9909	**37.837**	**0.9952**
2	26.166	0.9228	22.493	0.7761	24.053	0.8224	37.601	0.9935	35.949	0.9902	**37.745**	**0.9944**
3	27.241	0.9330	23.398	0.7950	24.884	0.8310	38.294	0.9933	37.364	0.9908	**39.396**	**0.9956**
4	35.385	0.9855	35.721	0.9697	31.202	0.9257	35.508	0.9857	42.158	0.9897	**43.941**	**0.9953**

Figure 8. Comparison of the effects of different demosaicking methods for the SateHaze1k dataset.

Figure 8 compares the subjective effects of different demosaicking algorithms on the SateHaze1k dataset. In scenario one, Bicubic, FlexISP, and ADMM algorithms exhibit varying degrees of color aliasing. Bicubic shows the most severe color aliasing, while FlexISP introduces pseudo-color artifacts in pixel form, resembling numerous pseudo-color noise points. Although ADMM shows relatively lighter aliasing compared to others, it still generates significant areas of pseudo-color that deviate noticeably from the ground truth, and while JDSR and DPIR do not produce extensive pseudo-artifacts, they exhibit mosaic-like artifacts in certain key structures, such as the blue block area on the far right, which are absent in the ground truth. In contrast, the proposed method effectively mitigates the aforementioned issues. In Scenario Two, we meticulously evaluated the visual outcomes at the pixel level. Bicubic interpolation, FlexISP, and ADMM methods displayed noticeable artifacts in the form of widespread false colors, undermining the accuracy and realism

of the reconstructed images. JDSR, on the other hand, struggled with color preservation, leading to a loss of chromatic information in the demosaicked images.
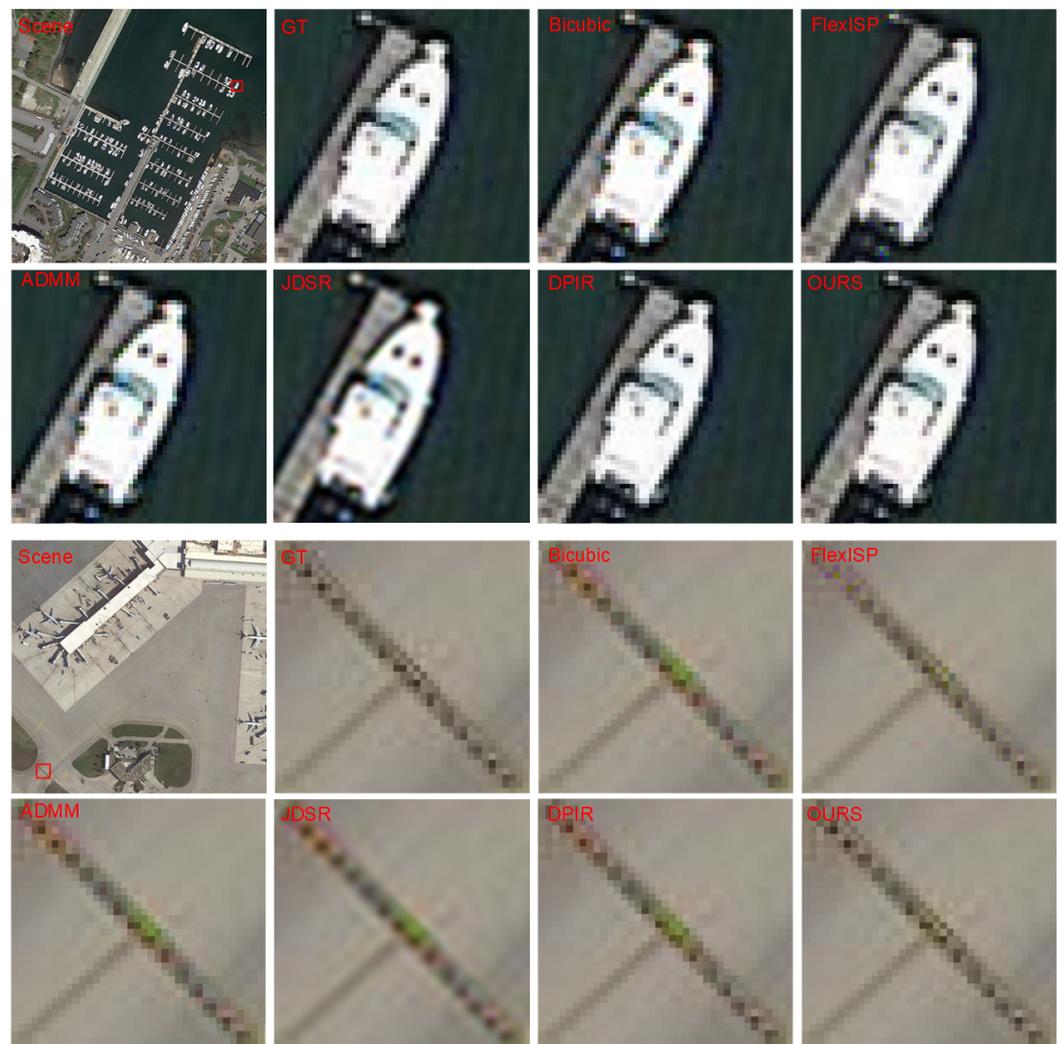

Figure 9. Comparison of the effects of different demosaicking methods for the DOTA-v1.0 dataset.

Notably, DPIR introduced extraneous colors that deviated from the ground truth, compromising the faithfulness of the color reproduction process. In contrast, our proposed method effectively restored the original pixel colors to a high degree of fidelity.

Figure 9 showcases the comparison results on the DOTA-v1.0 dataset. In the first scenario featuring a yacht, Bicubic and ADMM exhibit noticeable pseudo-color artifacts, FlexISP and DPIR suffer from color loss, and JDSR shows excessive smoothing. In comparison, the proposed algorithm closely aligns with the ground truth. In the second scenario depicting lines on an airport runway, Bicubic, FlexISP, ADMM, DPIR, and JDSR all demonstrate severe color aliasing. In contrast, the proposed algorithm effectively restores the colors and details of the lines on the runway. Figures 8 and 9 visually depict the comparative results of different demosaicking algorithms on the remote sensing images, highlighting the superior performance of our proposed method in preserving image details, reducing artifacts, and enhancing overall image quality. These findings demonstrate the efficacy and robustness of our proposed algorithm in handling demosaicking tasks on remote sensing datasets.

5. Discussion

In this section, we devised a series of experiments aimed at investigating the design and efficacy of two pivotal components: (1) the discriminant-enhanced learning module

and (2) the global attention mechanism. Through the implementation of ablation studies, we scrutinized the individual contributions of these components. To ensure the rigor and impartiality of our comparisons, all networks underwent training in an identical environment, facilitating a fair assessment. Consistency was maintained by utilizing the same training dataset and preprocessing operations across all experiments, and evaluations were conducted on the same datasets as above. By subjecting the networks to identical testing conditions, we sought to provide a robust evaluation of the specific impacts of the discriminator-enhanced learning module and the global attention mechanism on the demosaicking task. Our meticulous experimental design enabled a nuanced analysis of the distinct roles played by these key components in enhancing demosaicking performance. By systematically dissecting the contributions of the discriminator-enhanced learning module and the global attention mechanism, we gained valuable insights into their respective effects on the quality and accuracy of color image reconstruction.

5.1. Ablation Studies of Discriminant-Enhanced Learning (DEL) Link

The proposed method incorporates a progressive discriminator strategy, featuring a discriminator-enhanced learning module within its model architecture. Leveraging exponential moving average techniques, the discriminator-enhanced learning module enhances the stability of the generative model, thereby contributing to more consistent model performance. Additionally, this module facilitates the construction of a pixel-wise map indicating the likelihood of artifacts, ensuring the accuracy of optimization directions. These capabilities collectively contribute to the achievement of high-precision demosaicking results.

To demonstrate the efficacy of the network architecture, we adopted SRGAN as the base model and evaluated the impact of integrating the discriminator-enhanced learning module into the proposed demosaicking model, as illustrated in Table 3. By deploying this module, we observed significant improvements in the demosaicking performance, highlighting its role in enhancing the overall quality and fidelity of the reconstructed images. The utilization of the discriminator-enhanced learning module not only enhances the stability of the generative model but also enables the network to effectively address artifacts and optimize the demosaicking process for superior results. This strategic integration underscores the importance of incorporating advanced techniques, such as discriminator-enhanced learning, to elevate the performance of demosaicking algorithms and achieve state-of-the-art results in image demosaicking tasks.

Table 3. Ablation studies of Discriminant-enhanced Learning Link (DEL). The "$-$" and "$+$" represent the model without the DEL and the model with the DEL, respectively.

DEL	McMaster		Kodak		Set5		SateHaze1k		DOTA-v1.0	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
$-$	30.245	0.9251	30.572	0.9233	33.377	0.9329	24.198	0.8903	34.658	0.9324
$+$	**33.756**	**0.9688**	**33.687**	**0.9775**	**37.006**	**0.9679**	**27.591**	**0.9322**	**38.522**	**0.9893**

5.2. Ablation Studies of Global Attention Mechanism

The global attention mechanism serves as a crucial component within the demosaicking network. In order to validate the effectiveness of the global attention mechanism employed in this study for demosaicking tasks, we conducted performance evaluations across four scenarios: (1) a network without any attention mechanism; (2) a network incorporating Convolutional Block Attention Module (CBAM) [65]; (3) a network integrating Coordinated Attention (COA) [66]; and (4) a network featuring the Global Attention Mechanism (GAM), which corresponds to the method proposed in this paper, as depicted in Table 4.

The experimental results presented in Table 4 demonstrate that while a network without attention mechanisms can roughly accomplish the task of image demosaicking, its performance significantly lags behind networks equipped with attention mechanisms. Specifically, the global attention mechanism typically outperforms other attention mechanisms in demosaicking networks. For instance, on the Kodak dataset, the PSNR metric with

GAM improved by nearly 3 dB compared to CBAM and by over 2 dB compared to COA. Therefore, our proposed network with GAM effectively enhances demosaicking capabilities, showcasing superior performance in image demosaicking tasks. The results underscore the pivotal role of the global attention mechanism in improving demosaicking quality and highlight its potential to elevate the overall performance of demosaicking networks.

Table 4. Ablation studies of different Attention Mechanism (AM). The "−" and "+" represent the model without the AM and the model with the AM, respectively.

AM	McMaster		Kodak		Set5		SateHaze1k		DOTA-v1.0	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
−	30.245	0.9251	30.572	0.9233	33.377	0.9329	24.198	0.8903	34.658	0.9324
+CBAM	31.263	0.9276	31.552	0.9328	34.293	0.9452	25.466	0.9075	35.231	0.9458
+COA	32.153	0.9399	32.488	0.9598	35.256	0.9547	26.296	0.9183	36.248	0.9506
+GAM(OURS)	**34.233**	**0.9568**	**34.526**	**0.9740**	**37.381**	**0.9755**	**27.989**	**0.9366**	**39.257**	**0.9778**

6. Conclusions

This study presents a novel end-to-end remote sensing image demosaicking approach that harnesses Generative Adversarial Networks based on a global attention mechanism within. By integrating a progressive discriminative framework, we effectively differentiate between visual artifacts and authentic details. Furthermore, the integration of global attention mechanisms within the generator further enhances spatial-channel information interaction, leading to superior performance in image demosaicking tasks. Our investigation includes a comprehensive evaluation of multiple prevailing attention mechanisms in the demosaicking task. Through rigorous experimentation on a widely adopted dataset, we conclusively determine the global attention mechanism as the most suitable approach for this specific application. The results of our study showcase that our proposed method significantly outperforms existing state-of-the-art techniques in both quantitative metrics and qualitative visual fidelity. Through extensive experimentation and evaluation, our approach emerges as a robust solution, offering substantial advancements in the field of remote sensing image demosaicking.

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

Data Availability Statement: The data were prepared and analyzed in this study.

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

References

1. Chang, L.; Tan, Y.P. Effective use of spatial and spectral correlations for color filter array demosaicking. *IEEE Trans. Consum. Electron.* **2004**, *50*, 355–365. [CrossRef]
2. Gunturk, B.; Glotzbach, J.; Altunbasak, Y.; Schafer, R.; Mersereau, R. Demosaicking: Color filter array interpolation. *IEEE Signal Process. Mag.* **2005**, *22*, 44–54. [CrossRef]
3. Cok, D.R. Signal Processing Method and Apparatus for Producing Interpolated Chrominance Values in a Sampled Color Image Signal. U.S. Patent 4,642,678, 10 February 1987.

4. Adams, J.E., Jr. Interactions between color plane interpolation and other image processing functions in electronic photography. In Proceedings of the Cameras and Systems for Electronic Photography and Scientific Imaging, San Jose, CA, USA, 8–9 February 1995; Anagnostopoulos, C.N., Lesser, M.P., Eds.; International Society for Optics and Photonics, SPIE: Bellingham, WA, USA, 1995; Volume 2416, pp. 144–151. [CrossRef]
5. Kiku, D.; Monno, Y.; Tanaka, M.; Okutomi, M. Residual interpolation for color image demosaicking. In Proceedings of the 2013 IEEE International Conference on Image Processing, Melbourne, Australia, 15–18 September 2013; pp. 2304–2308. [CrossRef]
6. Kiku, D.; Monno, Y.; Tanaka, M.; Okutomi, M. Minimized-Laplacian residual interpolation for color image demosaicking. In *Digital Photography X*; Sampat, N., Tezaur, R., Battiato, S., Fowler, B.A., Eds.; International Society for Optics and Photonics, SPIE: Bellingham, WA, USA, 2014; Volume 9023, p. 90230L. [CrossRef]
7. Monno, Y.; Kiku, D.; Tanaka, M.; Okutomi, M. Adaptive residual interpolation for color image demosaicking. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; pp. 3861–3865. [CrossRef]
8. Hibbard, R.H. Apparatus and Method for Adaptively Interpolating a Full Color Image Utilizing Luminance Gradients. U.S. Patent 5,382,976A, 29 June 1993.
9. Adams, J. Design of practical color filter array interpolation algorithms for digital cameras .2. In Proceedings of the 1998 International Conference on Image Processing, ICIP98 (Cat. No.98CB36269), Chicago, IL, USA, 7 October 1998; Volume 1, pp. 488–492. [CrossRef]
10. Kakarala, R.; Baharav, Z. Adaptive demosaicing with the principal vector method. *IEEE Trans. Consum. Electron.* **2002**, *48*, 932–937. [CrossRef]
11. Buades, A.; Coll, B.; Morel, J.M.; Sbert, C. Self-Similarity Driven Color Demosaicking. *IEEE Trans. Image Process.* **2009**, *18*, 1192–1202. [CrossRef]
12. Lien, C.Y.; Yang, F.J.; Chen, P.Y.; Fang, Y.W. Efficient VLSI Architecture for Edge-Oriented Demosaicking. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *28*, 2038–2047. [CrossRef]
13. Kim, Y.; Jeong, J. Four-Direction Residual Interpolation for Demosaicking. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *26*, 881–890. [CrossRef]
14. Yang, X.; Zhou, W.; Li, H. MCFD: A Hardware-Efficient Noniterative Multicue Fusion Demosaicing Algorithm. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *31*, 3575–3589. [CrossRef]
15. Chen, X.; He, L.; Jeon, G.; Jeong, J. Multidirectional Weighted Interpolation and Refinement Method for Bayer Pattern CFA Demosaicking. *IEEE Trans. Circuits Syst. Video Technol.* **2015**, *25*, 1271–1282. [CrossRef]
16. Yang, X.; Zhou, W.; Li, H. Hardware-Oriented Shallow Joint Demosaicing and Denoising. In Proceedings of the 2022 IEEE International Conference on Image Processing (ICIP), Bordeaux, France, 16–19 October 2022; pp. 801–805. [CrossRef]
17. Mairal, J.; Elad, M.; Sapiro, G. Sparse Representation for Color Image Restoration. *IEEE Trans. Image Process.* **2008**, *17*, 53–69. [CrossRef]
18. Mairal, J.; Bach, F.; Ponce, J.; Sapiro, G.; Zisserman, A. Non-local sparse models for image restoration. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 29 September–2 October 2009; pp. 2272–2279. [CrossRef]
19. Ye, W.; Ma, K.K. Color Image Demosaicing Using Iterative Residual Interpolation. *IEEE Trans. Image Process.* **2015**, *24*, 5879–5891. [CrossRef]
20. Khashabi, D.; Nowozin, S.; Jancsary, J.; Fitzgibbon, A.W. Joint Demosaicing and Denoising via Learned Nonparametric Random Fields. *IEEE Trans. Image Process.* **2014**, *23*, 4968–4981. [CrossRef]
21. Klatzer, T.; Hammernik, K.; Knobelreiter, P.; Pock, T. Learning joint demosaicing and denoising based on sequential energy minimization. In Proceedings of the 2016 IEEE International Conference on Computational Photography (ICCP), Evanston, IL, USA, 13–15 May 2016; pp. 1–11. [CrossRef]
22. Kokkinos, F.; Lefkimmiatis, S. Deep Image Demosaicking Using a Cascade of Convolutional Residual Denoising Networks. In Proceedings of the Computer Vision—ECCV 2018, Munich, Germany, 8–14 September 2018; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Springer: Cham, Switzerland, 2018; pp. 317–333.
23. Kokkinos, F.; Lefkimmiatis, S. Iterative Joint Image Demosaicking and Denoising Using a Residual Denoising Network. *IEEE Trans. Image Process.* **2019**, *28*, 4177–4188. [CrossRef]
24. Syu, N.S.; Chen, Y.S.; Chuang, Y.Y. Learning Deep Convolutional Networks for Demosaicing. *arXiv* **2018**, arXiv:1802.03769.
25. Xu, Y.; Liu, Z.; Wu, X.; Chen, W.; Wen, C.; Li, Z. Deep Joint Demosaicing and High Dynamic Range Imaging Within a Single Shot. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 4255–4270. [CrossRef]
26. Chang, K.; Li, H.; Tan, Y.; Ding, P.L.K.; Li, B. A Two-Stage Convolutional Neural Network for Joint Demosaicking and Super-Resolution. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 4238–4254. [CrossRef]
27. Guan, J.; Lai, R.; Lu, Y.; Li, Y.; Li, H.; Feng, L.; Yang, Y.; Gu, L. Memory-Efficient Deformable Convolution Based Joint Denoising and Demosaicing for UHD Images. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 7346–7358. [CrossRef]
28. Wang, Y.; Yin, S.; Zhu, S.; Ma, Z.; Xiong, R.; Zeng, B. NTSDCN: New Three-Stage Deep Convolutional Image Demosaicking Network. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *31*, 3725–3729. [CrossRef]
29. Niu, G. Frequency Decomposition Network for Fast Joint Image Demosaic, Denoising and Super-Resolution. In Proceedings of the 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 6–8 January 2023; pp. 571–574. [CrossRef]

30. Vinod; Prasad, K.S.; Prasad, T.J. Deep Learning Approach for Image Denoising and Image Demosaicing. *Int. J. Comput. Appl.* **2017**, *168*, 18–26.

31. Sakamoto, T.; Nakanishi, C.; Hase, T. Software pixel interpolation for digital still cameras suitable for a 32-bit MCU. *IEEE Trans. Consum. Electron.* **1998**, *44*, 1342–1352. [CrossRef]

32. Hua, L.; Xie, L.; Chen, H. A color interpolation algorithm for Bayer pattern digital cameras based on green components and color difference space. In Proceedings of the 2010 IEEE International Conference on Progress in Informatics and Computing, Shanghai, China, 10–12 December 2010; Volume 2, pp. 791–795. [CrossRef]

33. Chung, K.L.; Hsu, T.C.; Huang, C.C. Joint Chroma Subsampling and Distortion-Minimization-Based Luma Modification for RGB Color Images With Application. *IEEE Trans. Image Process.* **2017**, *26*, 4626–4638. [CrossRef]

34. Sher, R.; Porat, M. CCD image demosaicing using localized correlations. In Proceedings of the 2007 15th European Signal Processing Conference, Poznan, Poland, 3–7 September 2007; pp. 1897–1901.

35. Menon, D.; Andriani, S.; Calvagno, G. Demosaicing With Directional Filtering and a posteriori Decision. *IEEE Trans. Image Process.* **2007**, *16*, 132–141. [CrossRef]

36. Nallaperumal, K.; Vinsley, S.; Christopher, S.; Selvakumar, R.K. A Novel Adaptive Weighted Color Interpolation Algorithm for Single Sensor Digital Camera Images. In Proceedings of the International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007), Sivakasi, India, 13–15 December 2007; Volume 3, pp. 477–481. [CrossRef]

37. Yang, B.; Wang, D. An Efficient Adaptive Interpolation for Bayer CFA Demosaicking. *Sens. Imaging* **2019**, *20*, 37. [CrossRef]

38. Sun, B.; Yuan, N.; Zhao, Z. A Hybrid Demosaicking Algorithm for Area Scan Industrial Camera Based on Fuzzy Edge Strength and Residual Interpolation. *IEEE Trans. Ind. Inform.* **2020**, *16*, 4038–4048. [CrossRef]

39. Fan, L.; Feng, G.; Ren, Y.; Wang, J. Color demosaicking via fully directional estimation. *SpringerPlus* **2016**, *5*, 1736. [CrossRef]

40. Su, D.; Willis, P. Demosaicing of color images using pixel level data-dependent triangulation. In Proceedings of the Theory and Practice of Computer Graphics, Birmingham, UK, 5 June 2003; pp. 16–23. [CrossRef]

41. Karch, B.K.; Hardie, R.C. Adaptive Wiener filter super-resolution of color filter array images. *Opt. Express* **2013**, *21*, 18820–18841. [CrossRef]

42. Jiaqi, W.; Taoyang, W.; Yufen, P.; Guo, Z. Bayer interpolation for video satellite images. *Remote. Sens. Nat. Resour.* **2019**, *31*, 51–58. [CrossRef]

43. Zhang, F.; Bai, C. Jointly Learning Spectral Sensitivity Functions and Demosaicking via Deep Networks. In Proceedings of the 2021 3rd International Conference on Advances in Computer Technology, Information Science and Communication (CTISC), Shanghai, China, 23–25 April 2021; pp. 404–411. [CrossRef]

44. Tang, J.; Li, J.; Tan, P. Demosaicing by Differentiable Deep Restoration. *Appl. Sci.* **2021**, *11*, 1649. [CrossRef]

45. Ignatov, A.; Van Gool, L.; Timofte, R. Replacing Mobile Camera ISP with a Single Deep Learning Model. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 13–19 June 2020; pp. 2275–2285. [CrossRef]

46. Sharif, S.M.A.; Naqvi, R.A.; Biswas, M. SAGAN: Adversarial Spatial-asymmetric Attention for Noisy Nona-Bayer Reconstruction. *arXiv* **2021**, arXiv:2110.08619.

47. Dong, W.; Yuan, M.; Li, X.; Shi, G. Joint Demosaicing and Denoising with Perceptual Optimization on a Generative Adversarial Network. *arXiv* **2018**, arXiv:1802.04723.

48. Huang, T.; Wu, F.F.; Dong, W.; Shi, G.; Li, X. Lightweight Deep Residue Learning for Joint Color Image Demosaicking and Denoising. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 127–132. [CrossRef]

49. Park, Y.; Lee, S.; Jeong, B.; Yoon, J. Joint Demosaicing and Denoising Based on a Variational Deep Image Prior Neural Network. *Sensors* **2020**, *20*, 2970. [CrossRef]

50. Khadidos, A.O.; Khadidos, A.O.; Khan, F.Q.; Tsaramirsis, G.; Ahmad, A. Bayer Image Demosaicking and Denoising Based on Specialized Networks Using Deep Learning. *Multimedia Syst.* **2021**, *27*, 807–819. [CrossRef]

51. Liang, J.; Zeng, H.; Zhang, L. Details or Artifacts: A Locally Discriminative Learning Approach to Realistic Image Super-Resolution. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, New Orleans, LA, USA, 18–24 June 2022. [CrossRef]

52. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 105–114. [CrossRef]

53. Liu, Y.; Shao, Z.; Hoffmann, N. Global Attention Mechanism: Retain Information to Enhance Channel-Spatial Interactions. *arXiv* **2021**, arXiv:2112.05561.

54. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.

55. Everingham, M.; Eslami, S.M. The Pascal Visual Object Classes Challenge: A Retrospective. *Int. J. Comput. Vision* **2015**, *111*, 98–136. [CrossRef]

56. Zhou, R.; Achanta, R.; Süsstrunk, S. Deep Residual Network for Joint Demosaicing and Super-Resolution. *Color Imaging Conf.* **2018**, *26*, 75–80. [CrossRef]

57. Xia, G.S.; Bai, X.; Ding, J.; Zhu, Z.; Belongie, S.; Luo, J.; Datcu, M.; Pelillo, M.; Zhang, L. DOTA: A Large-Scale Dataset for Object Detection in Aerial Images. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3974–3983. [CrossRef]

58. Huang, B.; Li, Z.; 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 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass Village, CO, USA, 1–5 March 2020; pp. 1795–1802. [CrossRef]

59. Heide, F.; Steinberger, M.; Tsai, Y.T.; Rouf, M.; Pająk, D.; Reddy, D.; Gallo, O.; Liu, J.; Heidrich, W.; Egiazarian, K.; et al. FlexISP: A Flexible Camera Image Processing Framework. *ACM Trans. Graph.* **2014**, *33*, 231. [CrossRef]

60. Tan, H.; Zeng, X.; Lai, S.; Liu, Y.; Zhang, M. Joint demosaicing and denoising of noisy bayer images with ADMM. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 2951–2955. [CrossRef]

61. Xu, X.; Ye, Y.; Li, X. Joint Demosaicing and Super-Resolution (JDSR): Network Design and Perceptual Optimization. *IEEE Trans. Comput. Imaging* **2020**, *6*, 968–980. [CrossRef]

62. Zhang, K.; Li, Y.; Zuo, W.; Zhang, L.; Van Gool, L.; Timofte, R. Plug-and-Play Image Restoration with Deep Denoiser Prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 6360–6376. [CrossRef]

63. Zhang, K.; Zuo, W.; Gu, S.; Zhang, L. Learning Deep CNN Denoiser Prior for Image Restoration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3929–3938.

64. Zerman, E.; Rana, A.; Smolic, A. Colornet—Estimating Colorfulness in Natural Images. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; pp. 3791–3795. [CrossRef]

65. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In Proceedings of the Computer Vision—ECCV, Munich, Germany, 8–14 September 2018; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Springer: Cham, Switzerland, 2018; pp. 3–19.

66. Hou, Q.; Zhou, D.; Feng, J. Coordinate Attention for Efficient Mobile Network Design. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 13708–13717. [CrossRef]