


Article

# Enhanced Solar Coronal Imaging: A GAN Approach with Fused Attention and Perceptual Quality Enhancement

Zhenhong Shang <sup>1,2,\*</sup>  and Ruiyao Li <sup>1</sup>

<sup>1</sup> Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China; 20212104075@stu.kust.edu.cn

<sup>2</sup> Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming 650500, China

\* Correspondence: szh@kust.edu.cn

**Abstract:** The activity of the solar corona has a significant impact on all aspects of human life. People typically use images obtained from astronomical telescopes to observe coronal activities, among which the Atmospheric Imaging Assembly (AIA) of the Solar Dynamics Observatory (SDO) is particularly widely used. However, due to resolution limitations, we have begun to study the application of generative adversarial network super-resolution techniques to enhance the image data quality for a clearer observation of the fine structures and dynamic processes in the solar atmosphere, which improves the prediction accuracy of solar activities. We aligned SDO/AIA images with images from the High-Resolution Coronal Imager (Hi-C) to create a dataset. This research proposes a new super-resolution method named SAFCSRGAN, which includes a spatial attention module that incorporates channel information, allowing the network model to better capture the corona's features. A Charbonnier loss function was introduced to enhance the perceptual quality of the super-resolution images. Compared to the original method using ESRGAN, our method achieved an 11.9% increase in Peak Signal-to-Noise Ratio (PSNR) and a 4.8% increase in Structural Similarity (SSIM). Additionally, we introduced two perceptual image quality assessment metrics, the Natural Image Quality Evaluator (NIQE) and Learned Perceptual Image Patch Similarity (LPIPS), which improved perceptual quality by 10.8% and 1.3%, respectively. Finally, our experiments demonstrated that our improved model surpasses other models in restoring the details of coronal images.

**Keywords:** super resolution; astronomical image; corona; GAN



**Citation:** Shang, Z.; Li, R. Enhanced Solar Coronal Imaging: A GAN Approach with Fused Attention and Perceptual Quality Enhancement. *Appl. Sci.* **2024**, *14*, 4054. <https://doi.org/10.3390/app14104054>

Academic Editor: Mario Gai

Received: 10 April 2024

Revised: 4 May 2024

Accepted: 7 May 2024

Published: 10 May 2024



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

## 1. Introduction

The sun, being the star closest to us, plays a significant role in our solar system, and phenomena such as coronal mass ejections (CMEs) during solar activities have widespread effects on the Earth. Intense solar activity could lead to catastrophic weather events on our planet, pose risks to space operations, and may even threaten human lives [1]. Therefore, monitoring and researching solar coronal activities is of great significance to humanity [2].

The Atmospheric Imaging Assembly (SDO/AIA) of the Solar Dynamics Observatory provides critical information for human observation of solar activities. The Coronal Mass Ejection on 13 June 2010, and its associated extreme ultraviolet wave, were studied using high-resolution core channel data of 193 Å and 211 Å from the Atmospheric Imaging Assembly [3,4] on board the Solar Dynamics Observatory [5]. The 171 Å filter of AIA is particularly suitable for observing plume-like structures and large coronal loops, features that are often found in the corona's quiescent areas and regions with weaker magnetic fields. The 193 Å filter of AIA is specifically appropriate for observing coronal structures within active regions, such as active coronal loops and solar flares. Yang et al. (2013) [6] used AIA's 193 and 211 channels to validate the fundamental physical mechanisms of the prevalent flare models, providing visual evidence of magnetic reconnection generating solar flares.

Xue et al. (2016) [7] demonstrated a new role in the solar eruptions' reconnection through AIA's 131, 171, and 304 bands: the release of magnetic torsion. Berghmans et al. (2021) [8] compared AIA with high-resolution EUV telescopes and found that "campfires" (Small-scale localized bright regions, also known as 'campfires,' appear in the quiet areas of the Sun, with lengths ranging from 400 to 4000 km and durations from 10 s to 200 s) are essentially coronal in nature, originating from the magnetic flux concentrations of the chromospheric network.

Higher resolution data also provide a more comprehensive perspective for the observation of solar activities. The High-resolution Coronal Imager (Hi-C) captured extremely clear coronal images in 193 angstrom and 172 angstrom wavelengths, respectively, in 2012 and 2018. These images had a resolution of at least 0.47 arcseconds, providing an unprecedented opportunity to study the mass and energy coupling between the chromosphere and the corona. Cirtain et al. (2013) [9] confirmed the existence of various heating mechanisms in the solar outer atmosphere, or corona, through the solar coronal images captured by Hi-C. Regnier et al. (2014) [10] found evidence of small-scale extreme ultraviolet (EUV) heating events at the bases of coronal loops in high-definition images. Kuznetsov (2015) [11] revealed detailed information that helps us understand the dynamics of the magnetic field, fine structures, and flare energy release. Barczynski (2017) [12] observed that the size, motion, and temporal evolution of coronal loop features are associated with photospheric motions, indicating that they are closely connected to the photospheric magnetic field. Williams et al. (2020) [13] discovered that current density is proportional to magnetic field strength and emphasized the importance of establishing a permanent solar observatory facility with high resolution similar to Hi-C. This motivates us to work on improving the quality of SDO/AIA images to match the resolution of Hi-C and to provide higher temporal resolution coronal images for solar imagery research. This has inspired us to endeavor to enhance the quality of SDO/AIA images to achieve the same resolution as that of Hi-C, and to furnish solar image research with coronal images of higher temporal resolution.

The resolution of observed solar images is constrained by the capabilities of the instruments. Even with the most advanced instruments available today, it is challenging to achieve the high resolution that theoretical solar physicists aspire to. Recently, various deep learning-based super-resolution models have been successfully applied to the processing of astronomical images. For instance, Rahman et al. (2020) [14] enhanced the clarity of SDO/HMI images using two deep learning models, and the results were consistent with those from the Hinode magnetograms. Yang et al. (2022) [15] provided solar EUV images with higher temporal resolution for space weather forecasting and solar physics research through super-resolution technology. Bi et al. (2023) [16] further confirmed the occurrence of magnetic reconnection within the braided structures of the corona by applying super-resolution processing to SDO/AIA images.

While the high-resolution coronal images provided by Hi-C have led to new discoveries for researchers, its intermittent observation times and limited coverage of solar regions pose certain limitations. In contrast, SDO/AIA can provide continuous and comprehensive coronal observation data. Hi-C captured solar coronal activities at wavelengths of 193 angstroms and 172 angstroms, while AIA continuously observes the entire sun at wavelengths of 193 angstroms and 171 angstroms. Due to the similarity of their central wavelength data, it is necessary to register and calibrate the images to achieve better super-resolution results. Image registration effectively eliminates heterogeneity in the original data, allowing us to use popular deep learning neural networks to learn the mapping relations of images, thereby enhancing the spatial resolution of AIA images to match the level of Hi-C.

In this paper, we utilized SDO/AIA images as low-resolution samples and Hi-C images as high-resolution samples. We constructed our dataset using SIFT image registration technology, simulating the scenario in actual scientific research where different devices capture solar astronomical images concurrently. Inspired by the ESRGAN [17] model, which offers decent image perceptual quality but performs poorly in pixel metrics, result-

ing in excessive smoothing effects during the super-resolution reconstruction of coronal images, there was a loss of coronal details, particularly in texture-rich regions. Additionally, the ESRGAN model employs a substantial number of dense connections, which leads to a large parameter size and, consequently, the model training requires extensive computational resources.

We proposed a new super-resolution method, SAFCSRGAN, and designed a hybrid channel-spatial attention mechanism that enhanced the model's capability to capture subtle coronal features. We also optimized the loss function, significantly improving the model's performance on key image quality assessment metrics and surpassing other models in perceptual quality. Compared to ESRGAN, our model's parameters were reduced by 21%. Experimental results confirmed the effectiveness of our proposed method.

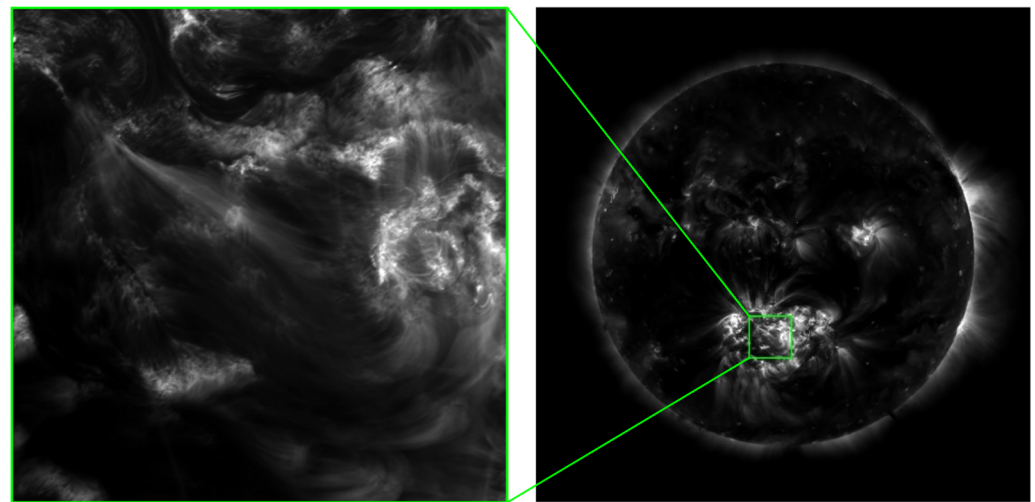
## 2. Data and Methods

### 2.1. Data Set

Different from the common practice in the field of super-resolution, where high-resolution (HR) images are downsampled to obtain low-resolution (LR) images, we applied deep learning mapping studies by aligning solar astronomical images captured at the same time but using different hardware facilities. This method provides LR and HR image comparisons that more closely match actual conditions, hence gathering data more suitable for real-world application scenarios, which is extremely beneficial for enhancing the generalization ability of the model. In the experiments presented in this paper, the LR images were sourced from the Atmospheric Imaging Assembly [3] (AIA) of the Solar Dynamics Observatory [4] (SDO), selecting images at two different wavelengths of 171 angstroms and 193 angstroms. These bands can effectively capture phenomena such as coronal mass ejections. The HR images are from NASA's High-Resolution Coronal imager (Hi-C) project, which includes all 114 high-definition coronal images from both Hi-C 1 and Hi-C 2.1. Hi-C 1 captured 36 coronal images at 193 angstroms from 16:52 to 16:55 on 11 July 2012, while Hi-C 2.1 captured 78 coronal images at 172 angstroms from 16:56 to 17:01 on 29 May 2018. To expand our dataset, we performed random cropping, horizontal flipping, rotation, and translation operations, ultimately selecting 440 image pairs as the training set. The resolution of HR images is  $2048 \times 2048$ , while the resolution of LR images is  $512 \times 512$ . In addition, we selected another 80 pairs of images as the validation set and the test set, with data from Hi-C 1 and Hi-C 2.1's two different wavebands divided in a 2:1 ratio.

### 2.2. Image Registration

Although the images collected by the Solar Dynamics Observatory (SDO/AIA) have a comprehensive field of view, their resolution is relatively low; in contrast, the High-Resolution Coronal Imager (Hi-C) captures images that cover only a partial area but have a higher resolution. In light of this, we need to register images from different sources that are captured at the same moment and of the same astronomical targets. We adopted the Scale-Invariant Feature Transform (SIFT) registration algorithm [18] to accomplish the registration work between high-resolution and low-resolution image pairs. This algorithm can extract key points (feature points) and their descriptors from the images, and these key points exhibit invariance to changes in image scale, rotation, and even brightness. For example, Figure 1a shows a high-definition coronal image from Hi-C; Figure 1b shows a lower-resolution image from SDO/AIA. After registration, the green box in Figure 1b indicates the low-resolution area corresponding to Figure 1a. Subsequently, we cropped the registered images to serve as the low-resolution (LR) images.



(a) Hi-C high resolution image      (b) SDO/AIA low resolution image

**Figure 1.** Coronal images are captured by different devices and the position of high-resolution images within low-resolution ones is determined through image registration.

### 2.3. Image Normalization

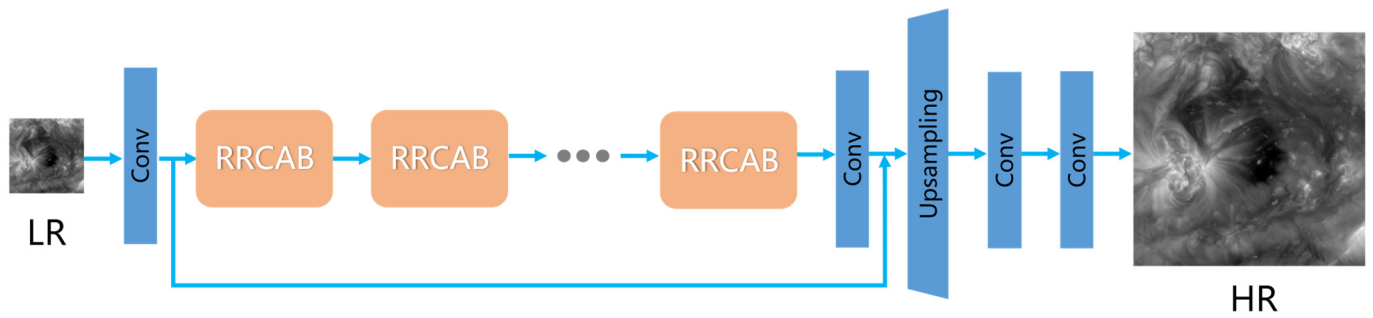
Due to the fact that SDO/AIA and Hi-C are two different instruments, the images they capture have different image attributes. In order to ensure that the dataset has the same scale, accelerate the training speed, prevent model gradient explosion, reduce the chance of the model falling into local optima, facilitate subsequent processing, and maintain numerical stability, we normalized the images. This standardization process adjusts the pixel values to a common scale, allowing the model to train more efficiently and effectively. We employ Z-score normalization to rescale the pixel values of the dataset images, resulting in images with a mean of 0 and a standard deviation of 1. For each pixel value  $I$  in the image, the normalized value  $I'$  is calculated as follows:

$$I' = 255 * \frac{I - \mu}{\sigma} \quad (1)$$

where  $\mu$  is the mean of all pixel values calculated from the images, and  $\sigma$  is the standard deviation.

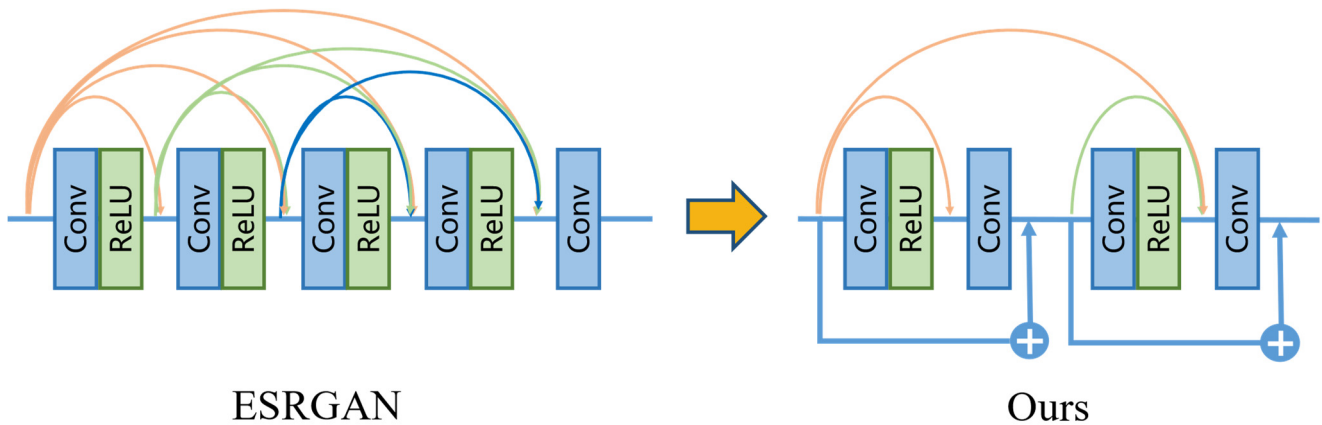
### 2.4. Network Structure of SAFCSRGAN

The generator network structure of the SAFCSRGAN (Spatial Attention of Fused Channels Super-Resolution GAN) proposed in this paper is shown in Figure 2. The generator mainly includes a shallow feature extraction module and a deep feature extraction and upsampling module [19]. The feature extraction module is responsible for learning the super-resolution features from low-resolution (LR) to high-resolution (HR) images, while the upsampling module is responsible for generating the final HR images based on the learned mapping relationship. The first layer of the network is a convolutional layer used for shallow feature extraction, and its output is also used as the input for both the upsampling module and deep feature extraction module. The core structure consists of a series of Residual in Residual Concatenation Attention Blocks (RRCAB), which can effectively achieve feature extraction and fusion, ensuring the training stability of the deep network. The residual connections within RRCAB help to maintain the flow of gradient information, thus avoiding the problem of gradient vanishing. After passing through a series of RRCABs, the feature maps are then reconstructed at a super-resolution by the upsampling module. The upsampling technique employed in this paper is PixelShuffle [20], also known as sub-pixel convolution.



**Figure 2.** The proposed SAFCSRGAN model structure includes a shallow feature extraction module, a deep feature extraction module, and an upsampling module.

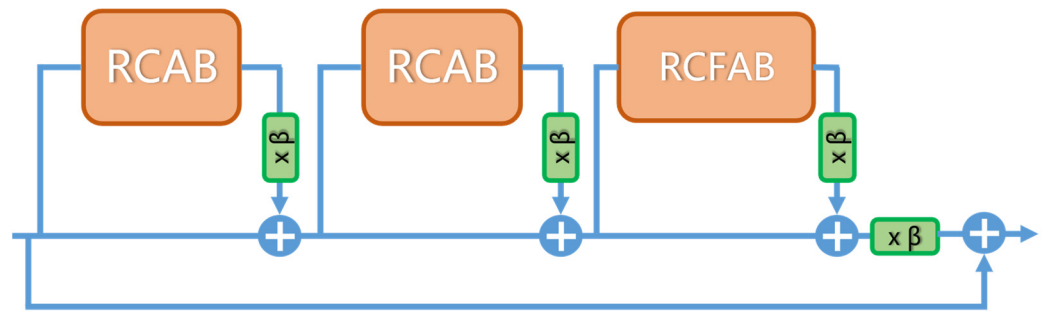
Our proposed method, compared to ESRGAN, removed some of the redundant dense connections. As shown in Figure 3, a basic block originally contained 9 connection operations. We believe that an excess of connections could potentially weaken model performance, as they might result in the preservation of insignificant features. Therefore, while retaining some of the necessary connections, we introduced residual connections. Experiments have confirmed that this improvement not only significantly reduced the number of model parameters but also enhanced model performance.



**Figure 3.** Remove the redundant dense connections from the original model’s basic blocks while introducing residual connections.

2.5. Residual in Residual Concatenation Attention Block (RRCAB)

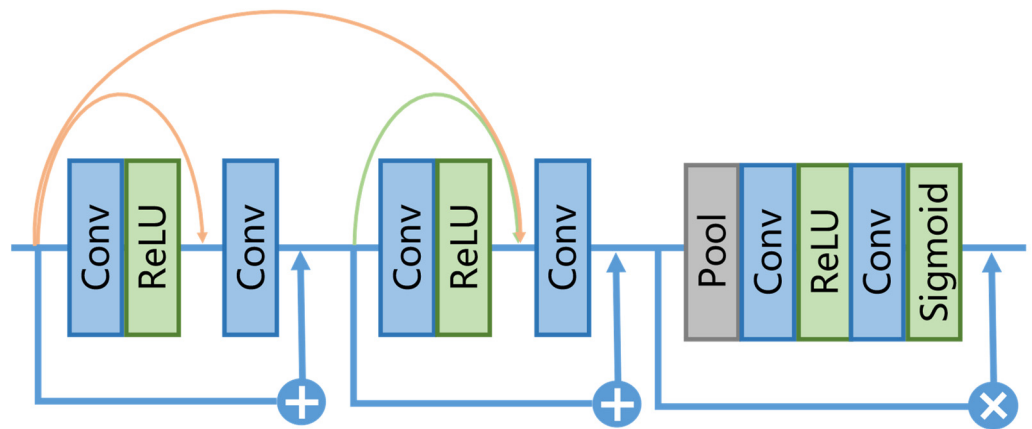
We propose a new model based on the existing one, which includes a Residual-in-Residual Concatenation Attention Block (RRCAB). As shown in Figure 4, the model structure contains two types of basic blocks: the Residual Concatenation Attention Block (RCAB) and the Residual Concatenation Fusion Attention Block (RCFAB) [21]. These two basic blocks mainly perform different feature extraction tasks, with the RCAB focusing on feature extraction among image channels, while the RCFAB emphasizes the extraction of spatial features of images and the fusion of channel and spatial features. These modules are connected in series through residual connections and introduce a scaling factor so that the network can learn the identity mapping and ensure the lossless transmission of information. Thus, even with a significant depth, the network can be trained effectively, improving the speed and accuracy of training.



**Figure 4.** Residual in Residual Concatenation Attention Block (RRCAB). Each RRCAB consists of two RCABs and one RCFAB. The former is responsible for extracting channel features, while the latter is tasked with extracting spatial features and fusing them with the channel features.

### 2.5.1. Residual Concatenation Attention Block (RCAB)

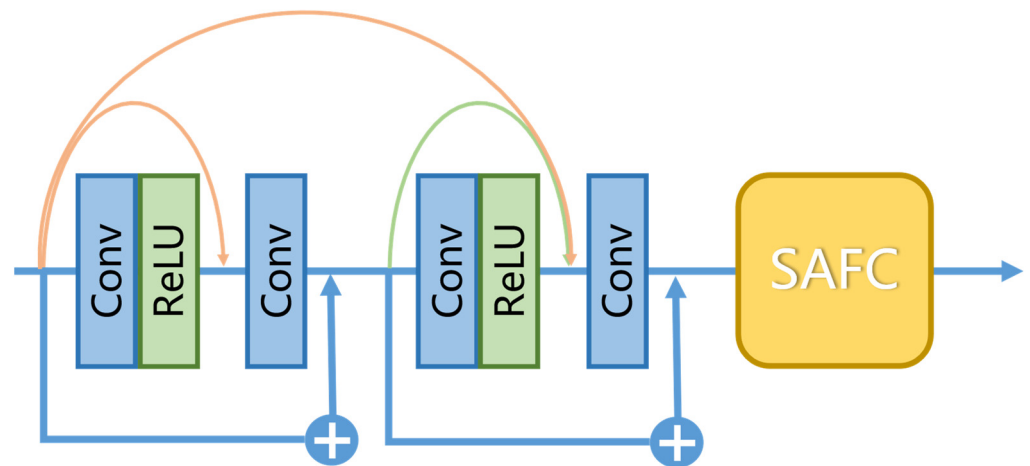
The Residual Concatenation Attention Block (RCAB) proposed in this paper improves upon the numerous dense connections used in the original model, as illustrated in Figure 5. Experiments indicate that dense connections consume a significant amount of computational resources, especially during forward propagation, because they retain the outputs of all previous layers, which in turn lowers network performance. To address this, we introduce the Residual Concatenation Attention Block, which, while maintaining some of the dense connections, incorporates a channel attention mechanism [22]. This mechanism enhances the model’s ability to recognize the importance of different channels by assigning different weights to them, allowing the network to focus more on features that are more important for the current task.



**Figure 5.** Residual Concatenation Attention Block (RCAB). By eliminating redundant dense connections and introducing residual connections and channel attention, the RCAB is constructed.

### 2.5.2. Residual Concatenation Fusion Attention Block (RCFAB)

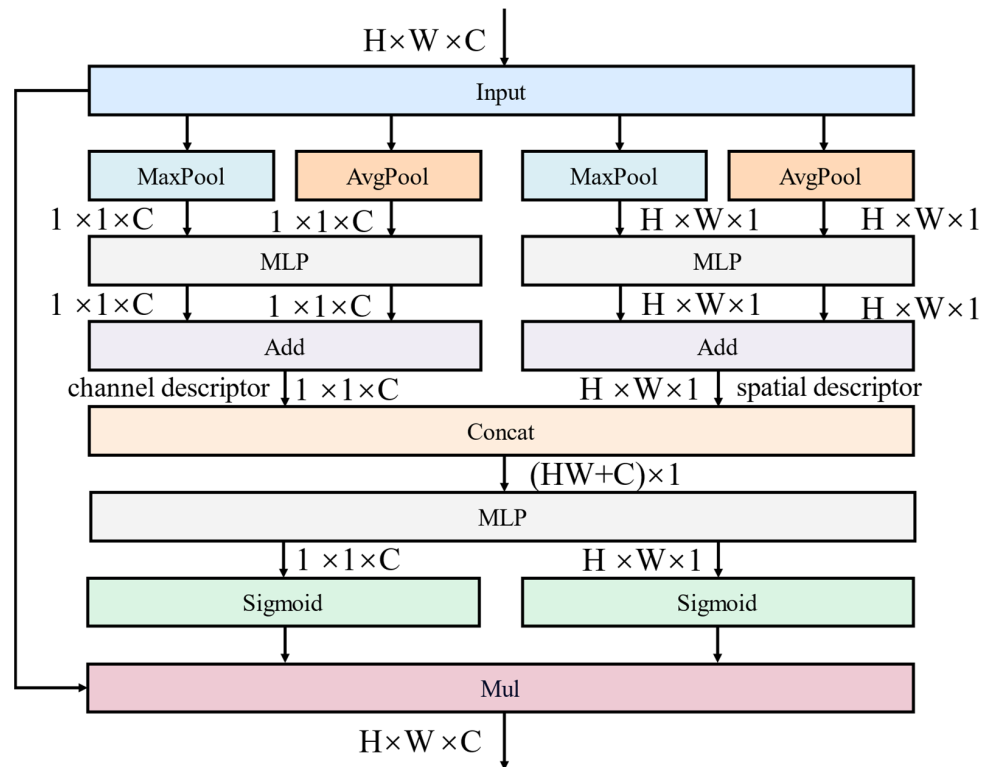
In another fundamental module, in the Residual Concatenation Fusion Attention Block (RCFAB), as shown in Figure 6, we have introduced a spatial attention mechanism that fuses channels. This enables the model to focus on the most critical spatial regions in the input data. By assigning different weights to different spatial positions, the model’s attention to key areas is enhanced. The introduction of the spatial attention mechanism greatly improves the model’s performance in super-resolution processing, especially when dealing with solar astronomical images that feature fine and dense coronal textures.



**Figure 6.** Residual Concatenation Fusion Attention Block (RCFAB). This block enhances the capability of capturing texture features by integrating a spatial attention of Fused Channels (SAFC).

2.5.3. Spatial Attention of Fused Channels (SAFC)

Figure 7 illustrates the spatial attention structure for the fusion channel that we propose. In this attention module, we hypothesize that the same information input, after being described through different channels and spatial representations, can exhibit diverse manifestations. We surmise that some feature information may be hidden between the channel and spatial descriptions. Consequently, we adopt a method that first isolates this feature information, then processes and integrates it. This separation–fusion operation not only retains the features captured by the previous network layer but also enhances the layer’s ability to learn weights that combine both channel and spatial features.



**Figure 7.** Spatial Attention of Fused Channels. This attention mechanism separates the same input and processes it through different descriptors to obtain spatial and channel features, which are then fused after processing.

## 2.6. Loss Function

To address the issue of excessive image smoothing and lack of sensory realism caused by the use of the L1 loss function in the original model, we have adopted the Charbonnier loss function [23] as a replacement. Lai et al. have demonstrated that compared to the L1 loss function, the Charbonnier loss function has a stronger capability to handle outliers, thereby being more effective in enhancing the performance of image super-resolution [24].

$$\text{Charbonnier}_{\text{Loss}} = \sqrt{x^2 + \epsilon^2} \quad (2)$$

wherein,  $x$  denotes the interpolation between the predicted value and the actual value, and  $\epsilon$  is a small constant to prevent the non-existence of the derivative when  $x = 0$ . In the model proposed in this paper,  $\epsilon$  is set to  $10^{-3}$ . Therefore, the total loss for the generator is as follows:

$$L_G = L_{\text{percep}} + \lambda L_G^{Ra} + \eta L_{\text{charbonnier}} \quad (3)$$

Perceptual loss  $L_{\text{percep}}$  is a type of loss function that constrains the features of network activations prior to activation layers. It is usually defined on the intermediate activation layers of a pre-trained deep network, which, however, may render the feature activations sparse. This phenomenon is particularly evident within deep learning frameworks. Furthermore, it could result in brightness inconsistencies between the super-resolution reconstructed images and the benchmark true images. To address these issues, this study has taken specific measures before the activation layers to ensure that the model can more accurately recover brightness and texture details. The definition of perceptual loss is as follows:

$$L_{\text{percep}} = \frac{1}{WHC} \sum_{x=1}^W \sum_{y=1}^H \sum_{z=1}^C \left( \Phi(SR)_{xyz} - \Phi(HR)_{xyz} \right)^2 \quad (4)$$

where  $W$ ,  $H$  and  $C$  respectively, represent the width, height, and number of channels of the feature map at the respective layer of the VGG network, and  $\Phi(SR)_{xyz}$  and  $\Phi(HR)_{xyz}$  denote the feature values at the coordinates  $((x, y, z))$  of the super-resolution image and high-resolution image, respectively.  $L_G^{Ra}$  is the adversarial loss, which is used to train the generator and discriminator in the generative adversarial network. The objective of the generator is to create high-resolution images that are realistic enough to “fool” the discriminator, while the discriminator attempts to distinguish between real and generated images. Since the adversarial loss occurs in tandem for the generator and the discriminator, the definition of adversarial losses for both is as follows:

$$L_G^{Ra} = -\mathbb{E}_{x_r} \left[ \log \left( 1 - D_{Ra} \left( x_r, x_f \right) \right) \right] - \mathbb{E}_{x_f} \left[ \log \left( D_{Ra} \left( x_f, x_r \right) \right) \right] \quad (5)$$

$$L_D^{Ra} = -\mathbb{E}_{x_r} \left[ \log \left( D_{Ra} \left( x_r, x_f \right) \right) \right] - \mathbb{E}_{x_f} \left[ \log \left( 1 - D_{Ra} \left( x_f, x_r \right) \right) \right] \quad (6)$$

wherein,  $x_f = G(x_i)$ , with  $x_i$  represents the input low-resolution (LR) image,  $D_{Ra}(x_r, x_f) = \sigma \left( C(x_r) - \mathbb{E}_{x_f} \left[ C(x_f) \right] \right)$ , and  $\mathbb{E}_{x_f}[\cdot]$  denote the operation of taking the average over all fake data in a mini-batch,  $\sigma$  is the sigmoid function [6], and  $C(x)$  represents the raw output of the discriminator.

## 3. Results and Analysis

### 3.1. Training Details

To increase efficiency and save resources, we set the batch size to 16 and cropped images in the dataset. The reason is that high-resolution images can increase processing time and resource consumption. We employed a uniform cropping method for both high-resolution (HR) images with 2K resolution and low-resolution (LR) images with 512 resolution. Specifically, each HR and LR image was uniformly cropped into 64 smaller segments. This processing method facilitates our model training.



We divided the training process into two main stages to optimize the performance of the super-resolution network model. In the initial stage, we use the L1 loss function to measure pixel-level differences, known as pixel loss, with the goal of cultivating a model that highly values the peak signal-to-noise ratio (PSNR). The learning rate for the initial model weights was set to  $1 \times 10^{-4}$ , employing an exponential decay strategy, where the learning rate is halved after every  $2 \times 10^{-5}$  mini-batch updates. This strategy helps in the gradual and fine-tuning of model parameters during the training process.

Subsequently, the model which is already well trained towards PSNR is used as the initial state of the generator within the GAN framework during the second stage. At this stage, the training of the generator adopts the Charbonnier composite loss function which replaces the L1 loss, with parameters  $\lambda = 5 \times 10^{-3}$  and  $\eta = 1 \times 10^{-2}$ . Specifically, the learning rate starts at  $1 \times 10^{-4}$  and is halved after reaching preset iteration thresholds (50 K, 100 K, 150 K, and 200 K). The purpose of this strategy is to balance efficiency with sufficient detail exploration in the learning process, avoiding too rapid a convergence to suboptimal solutions. For the optimizer, we use the Adam optimizer with  $\beta_1=0.9$  and  $\beta_2=0.999$ . The generator and discriminator are trained alternately until the model converges, with the generator containing 23 RRCAB blocks.

### 3.2. Evaluation Index

In this paper, we use four metrics to evaluate the super-resolution quality of solar images: PSNR, SSIM, NIQE [25], and LPIPS [26]. PSNR and SSIM are two metrics primarily used to quantify the similarity between super-resolution images and original images at the pixel level and structural level. NIQE and LPIPS, on the other hand, are used to measure the perceptual quality of images, reflecting their visual preference.

Given a ground truth (GT) image  $I$  with  $N$  pixels and a super-resolved image  $\hat{I}$ , the formula for calculating PSNR between  $I$  and  $\hat{I}$  is as follows, where a higher value of this metric indicates a smaller pixel differences between the images:

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{L^2}{\frac{1}{N} \sum_{i=1}^N (I(i) - \hat{I}(i))^2} \right) \quad (7)$$

The formula for calculating SSIM is as follows, where a higher value of this metric indicates smaller structural differences between the images.

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (8)$$

where  $\mu_x, \mu_y, \sigma_x, \sigma_y$ , and  $\sigma_{xy}$  are the means, variances, and covariance of the GT image and the super-resolved image, respectively.

The formula for calculating NIQE is as follows, where a smaller value of this metric indicates a higher natural perceptual quality of the image:

$$\text{NIQE} = \sqrt{(v_1 - v_2)^T \left( \frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (v_1 - v_2)} \quad (9)$$

where  $v_1$  and  $\Sigma_1$  are the mean vector and covariance matrix obtained from a set of sharp images, and  $v_2, \Sigma_2$  are the mean vector and covariance matrix obtained from the super-resolved image.

The calculation formula for LPIPS is as follows: the image is fed into a neural network, features are extracted from the L layer and then unit-normalized across the channel dimension. A vector  $w_1$  is used to scale the number of active channels, and the L2 distance is subsequently calculated. The result is averaged spatially and summed across the channels.

A smaller value of this metric indicates a perceptual quality of the image that is closer to the GT image.

$$\text{LPIPS} = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \|w_l \odot (\hat{y}_{hw}^l - \hat{y}_{0\ hw}^l)\|_2^2 \quad (10)$$

### 3.3. Experimental Results

On the test set, we compared our final model with the current state-of-the-art convolutional neural network-based super-resolution networks, including SRCNN [27], EDSR [28], RCAN [29], SRGAN [30], ESRGAN, and MSRResNet [31]. As shown in Table 1, our experimental results surpassed other models on multiple image quality metrics, particularly standing out in the aspect of perceptual quality of images. Our super-resolution experimental results are more in line with human visual perception. Compared with the baseline model ESRGAN, our model achieved an 11.92% improvement in the PSNR index and a 4.8% increase in the SSIM index. On 2 perceptual quality assessment metrics, NIQE and LPIPS were improved by 10.78% and 1.25%, respectively. Moreover, our model reduced the number of parameters by 21%, achieving higher efficiency. The results show that GAN-type networks tend to outperform other networks in terms of perceptual quality, although other types of models might achieve better results in PSNR and SSIM, their perceived quality usually does not match that of GAN-type networks.

**Table 1.** Comparison of image quality assessment metrics between ours and other models. In the super-resolution processing effect on solar corona images, ours outperforms other models in both pixel-level and perceptual metrics.

Method	PSNR↑	SSIM↑	NIQE↓	LPIPS↓
Bicubic	29.0103	0.7328	10.8501	0.3964
ESRGAN (Baseline)	32.9796	0.8762	5.9583	0.1835
SRCNN	30.0458	0.7891	9.9584	0.3781
EDSR	36.9069	0.9132	9.8624	0.3538
RCAN	36.4203	0.9030	9.7381	0.3562
SRGAN	28.5083	0.7796	7.2463	0.2954
MSRResNet	36.8074	0.9114	9.5365	0.3514
Ours	36.9125	0.9183	5.3158	0.1812

We compared the performance of several super-resolution models on the details of coronal images, as presented in Figure 8. The EDSR model achieved excellent results on the PSNR metric; however, its performance on the LPIPS index was not as good as the model we proposed. As can be seen from Figure 8c, the MSRResNet model produced severe anomalous dark spots in the coronal images after super-resolution processing. Similarly, the super-resolution image generated by the EDSR model exhibited haze-like artifacts. Despite this, both models scored higher on the SSIM index than our model. This presents a clear inconsistency with people's intuitive understanding of image quality.

### 3.4. Ablation Experiment

We have made critical improvements to the ESRGAN network. Firstly, an attention mechanism was introduced, which is particularly important for coronal images because it allows the model to capture coronal features more effectively. After the sixth RRCAB module in the network, we implemented feature activation visualization. Compared to the original model, our improved model demonstrates superior performance in capturing overall solar coronal features. As shown in Figure 9, the feature activation maps indicate that the introduction of the spatial attention mechanism significantly enhances feature recognition capability. Based on this, the model we propose not only further extends the range of capturing coronal features but also captures more coronal feature information that the original model failed to recognize.

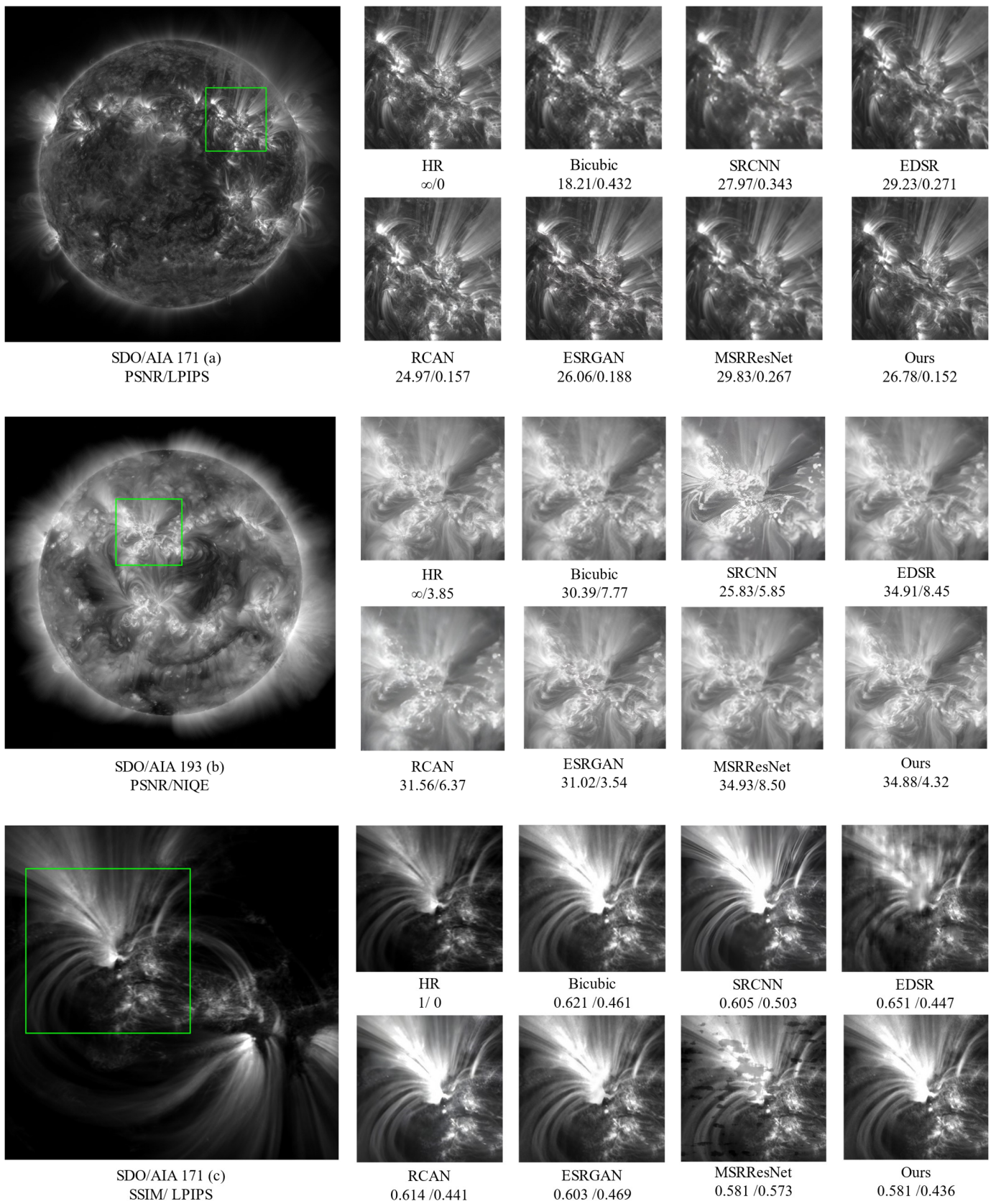
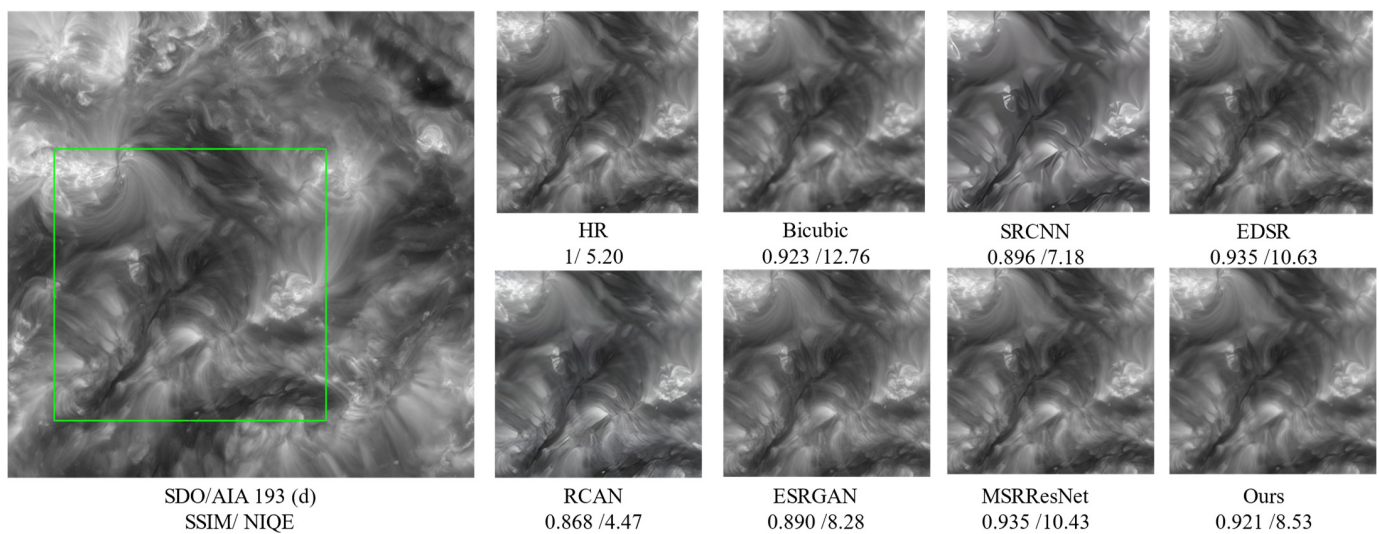
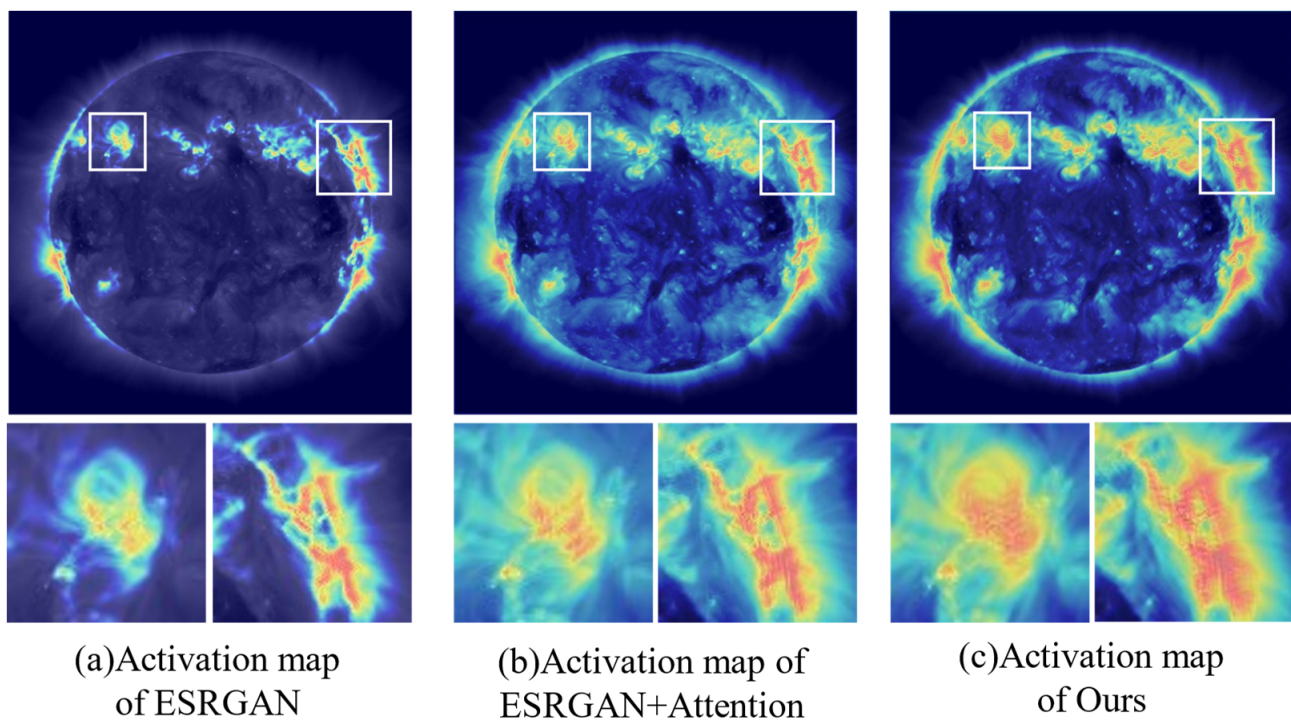


Figure 8. Cont.



**Figure 8.** Comparison of super-resolution details among different models. In (a,c), our method achieves the best performance in terms of the LPIPS metric, and in (c), our method does not have artifacts. In (b,d), our method strikes a good balance between pixel-level metrics and image perception metrics, with near-optimal values on pixel image metrics while also surpassing other models in terms of perceptual quality metrics.



**Figure 9.** Compare the feature activation maps of our model with the baseline model that incorporates the attention mechanism. Our model exhibits a broader scope and stronger focus on the features of corona images.

Secondly, we redesigned the architecture of the basic blocks while retaining some dense connections. We believe that a moderate amount of dense connections can enhance the utilization of shallow features, but the originally excessive dense connections led to the production of redundant features, which reduced the efficiency of the model's feature extraction. The improved model not only achieved better performance in terms of image super-resolution metrics, but also the number of model parameters was reduced by 21%.

As shown in Table 2, we removed some of the basic blocks with dense connections as part of the model's main body. As a result, the number of parameters was reduced to 13.02 M, yet the performance of the model was enhanced. This result confirms our hypothesis: an excess of dense connections, while increasing the amount of learnable parameters of the model actually limited the model's performance. In experiments, we observed that whether it was channel attention or spatial attention across merging channels, both significantly improved the model's metrics.

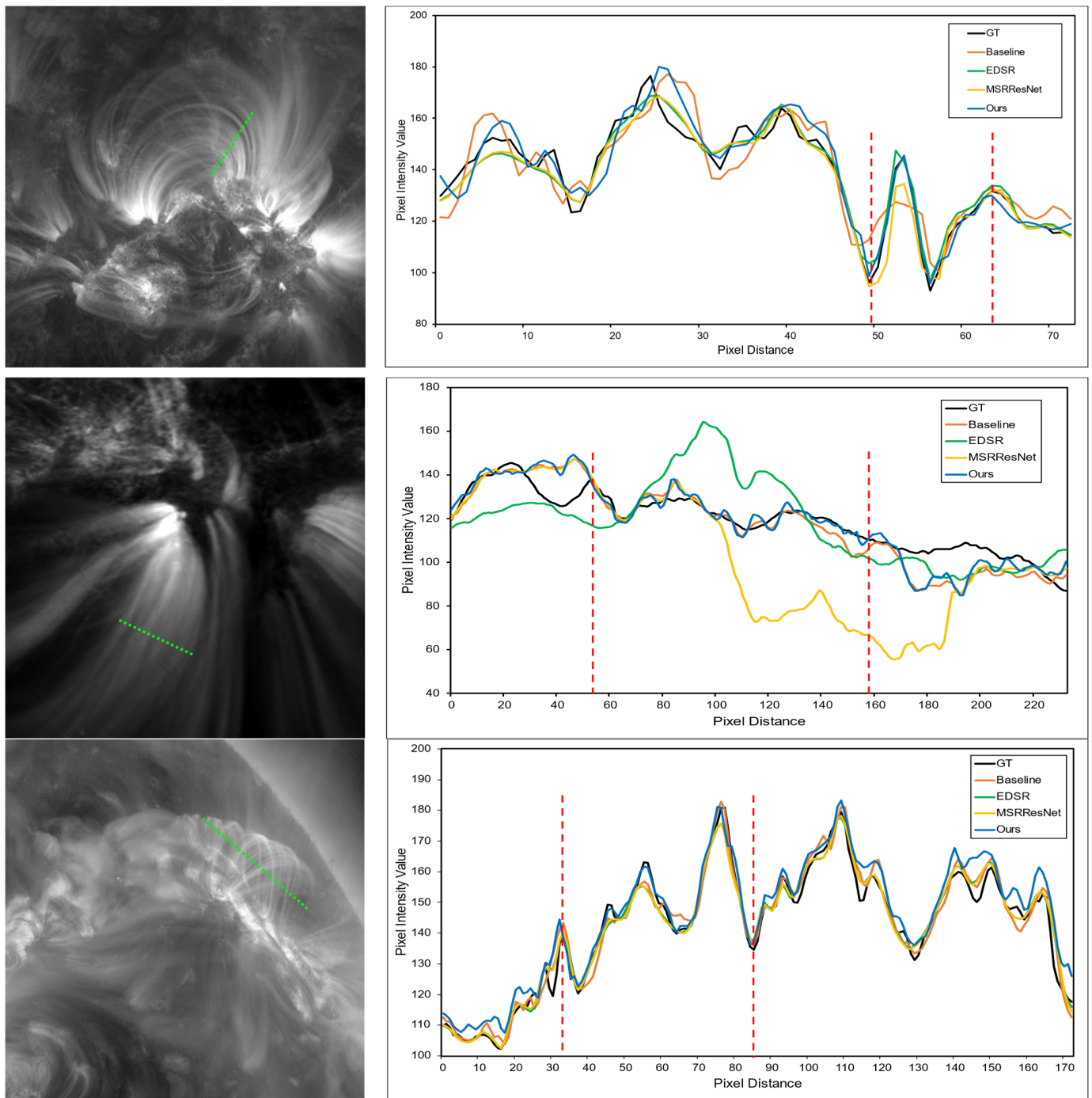
**Table 2.** The impact of different methods on the model.

Method	Choice				
	w	w/o	w/o	w/o	w/o
Dense concat	w	w/o	w/o	w/o	w/o
RCAB	w/o	w/o	w	w/o	w
RCFAB	w/o	w/o	w/o	w	w
PSNR	32.97	33.23	35.78	35.12	36.91
SSIM	0.87	0.84	0.89	0.88	0.91
NIQE	5.95	5.92	5.54	5.76	5.31
LPIPS	0.1835	0.1833	0.1821	0.1823	0.1812
Parameters (M)	16.69	13.02	13.11	13.06	13.16

#### 4. Discussion

We evaluated the AIA images that had undergone super-resolution processing and achieved results surpassing other models in terms of image quality metrics. To substantiate whether our approach could more accurately approximate Hi-C images compared to other models, we designed and conducted two experimental analyses for comparison. As shown in Figure 10, the profile lines [13] taken from three solar coronal images, after being processed by our model, perform closer to the original images in terms of image pixel intensity restoration. Especially in handling the pixels where the intensity rapidly decreases and then increases in the processed images, our model demonstrates superior recovery ability.

To validate the superiority of our model compared to the baseline model in restoring solar astronomical texture features, we conducted a difference comparison experiment. Specifically, we set three different thresholds to measure the acceptable differences between pixels in the super-resolved images and the original images. Pixels that exceeded this threshold were considered to have significant differences and were marked in green in the images. To make the results easier to observe, we also label the four surrounding pixel coordinates—up, down, left, and right of the discrepancy points—as green. At the same time, we calculated the percentage of green-marked pixels in the total image pixels, namely the rate of difference. Figure 11 shows the experimental results; our model has a significantly lower rate of difference compared to traditional models, with a noticeable reduction in green-marked difference points. This indicates that our model is more accurate in restoring the texture features of the solar surface compared to the baseline model, and the frequency of erroneous pixels is lower.



**Figure 10.** Compare the pixel intensity values along the line in coronal images for different models. Ours demonstrates better performance by more closely fitting the curves of the GT image in the annotation of pixel intensity curves.

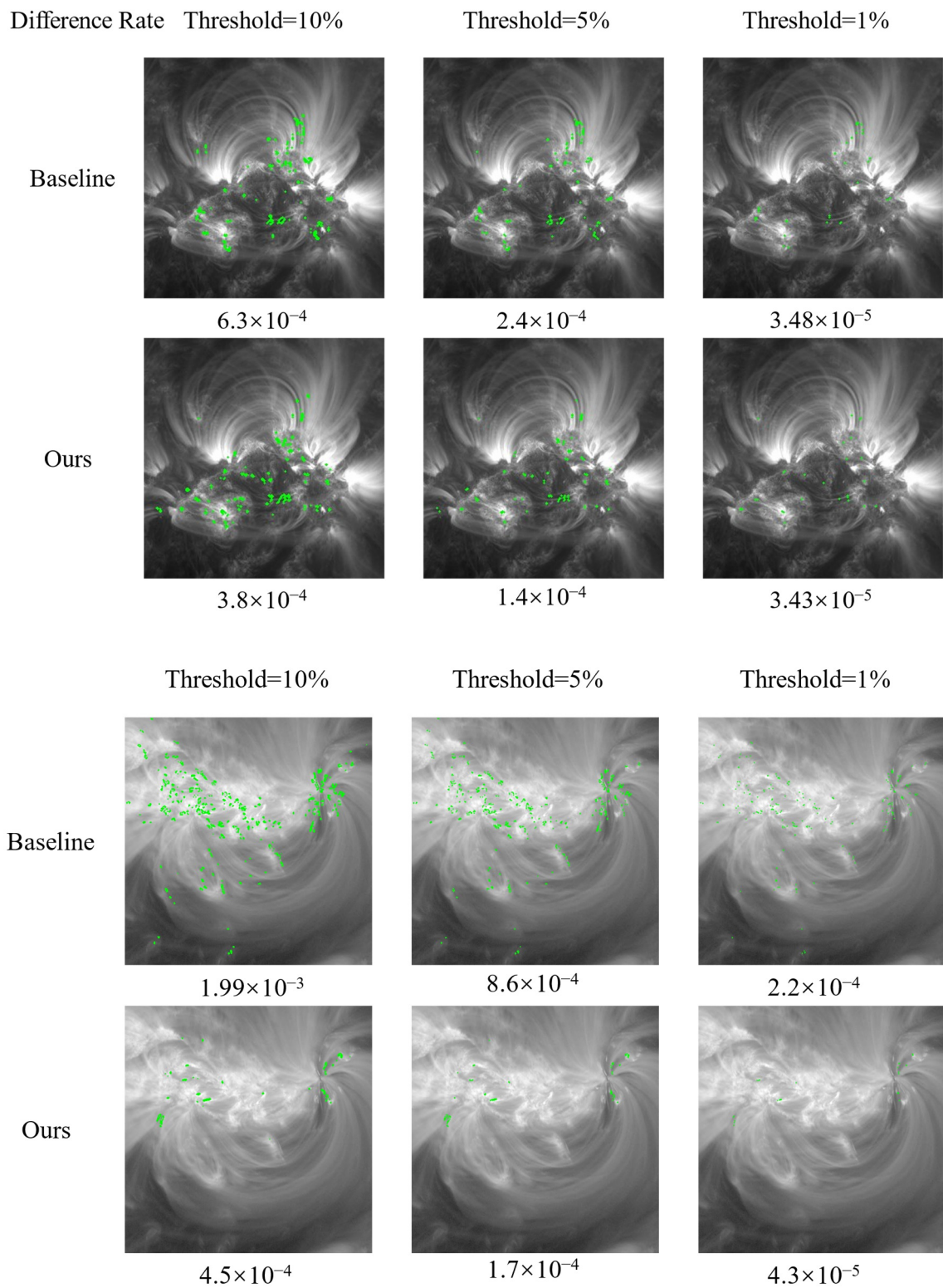
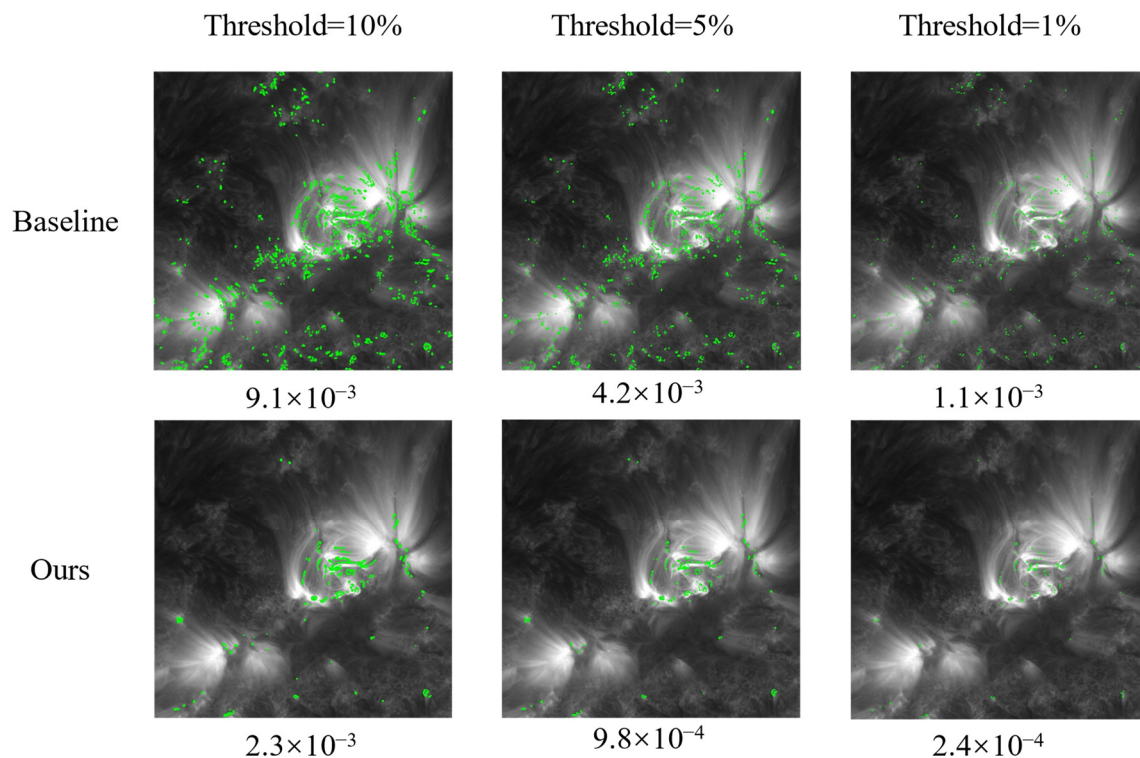


Figure 11. Cont.



**Figure 11.** Set the proportion of different pixels to total image pixels under different difference thresholds. Compared to the baseline model, ours has fewer points of discrepancy, indicating that the processed corona images are more similar to the original images.

## 5. Conclusions

In this study, we used SDO/AIA images as low-resolution inputs and Hi-C images as high-resolution targets to prepare datasets through image registration technology, creating paired images that are closer to real-world application scenarios. We improved the ESRGAN algorithm and proposed a new model named SAFCSRGAN, which we applied to the super-resolution reconstruction task of SDO/AIA images. Although the number of parameters in the model was reduced, the performance of our model has been enhanced. Experimental evidence shows that our improved model outperforms other common convolutional neural network super-resolution models in terms of performance. On image quality assessment metrics, our model has shown improvements in PSNR and SSIM indices compared to the original model. Additionally, on image perceptual quality assessment indices such as NIQE and LPIPS, our model also displayed superior performance compared to other models. Particularly in the application of super-resolution solar astronomical images, our model demonstrates better performance in reconstructing the texture details of images and produces a lower pixel difference rate compared to other models.

Super-resolution (SR) technology faces certain challenges when dealing with solar images. Our research indicates that deep learning-based models can be effectively used for super-resolution reconstruction of astronomical images, as they are able to recover the original texture details of the images well. Super-resolution images help us to study small-scale details in astronomical images, such as coronal morphological features and solar sunspot structures. However, limited by the scale of the dataset, the variety of images, and the performance of the models, there might be biases in the super-resolution effect for certain types of coronal images. In the future, we could consider expanding the dataset using images acquired by other coronal telescopes to enhance the model's generalization capability.

We would like to express our gratitude to the NASA SDO/AIA scientific team and the High-Resolution Coronal Imager (Hi-C) instrument team for providing the data.



**Author Contributions:** Conceptualization, Z.S.; funding acquisition, Z.S.; supervision, Z.S.; writing—review and editing, Z.S.; methodology, R.L.; software, R.L.; visualization, R.L.; writing—original draft, R.L.; validation, R.L.; formal analysis, R.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China, grant number 12063002.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All our research data are from public data. Hi-C data comes from [https://hic.msfc.nasa.gov/data\\_products.html](https://hic.msfc.nasa.gov/data_products.html), and SDO/AIA data comes from <http://jsoc.stanford.edu/data/aia/images/>. If you need the training set and test set data of this study, you can ask the corresponding author for it.

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

## References

1. Lockwood, M. Solar Influence on Global and Regional Climates. *Surv. Geophys.* **2012**, *33*, 503–534. [[CrossRef](#)]
2. Schwenn, R. Space Weather: The Solar Perspective. *Living Rev. Sol. Phys.* **2006**, *3*, 2. [[CrossRef](#)]
3. Lemen, J.R.; Title, A.M.; Akin, D.J.; Boerner, P.F.; Chou, C.; Drake, J.F.; Duncan, D.W.; Edwards, C.G.; Friedlaender, F.M.; Heyman, G.F.; et al. The Atmospheric Imaging Assembly (AIA) on the Solar Dynamics Observatory (SDO). *Sol. Phys.* **2012**, *275*, 17–40. [[CrossRef](#)]
4. Pesnell, W.D.; Thompson, B.J.; Chamberlin, P.C. The Solar Dynamics Observatory (SDO). In *The Solar Dynamics Observatory*; Chamberlin, P., Pesnell, W.D., Thompson, B., Eds.; Springer: New York, NY, USA, 2012; pp. 3–15, ISBN 978-1-4614-3673-7.
5. Ma, S.; Raymond, J.C.; Golub, L.; Lin, J.; Chen, H.; Grigis, P.; Testa, P.; Long, D. Observations and interpretation of a low coronal shock wave observed in the EUV by the SDO/AIA. *Astrophys. J.* **2011**, *738*, 160. [[CrossRef](#)]
6. Yang, R.; Wang, W. Comparison of Super-Resolution Reconstruction Algorithms Based on Texture Feature Classification. In Proceedings of the 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC), Xiamen, China, 5–7 July 2019; pp. 306–310.
7. Xue, Z.; Yan, X.; Cheng, X.; Yang, L.; Su, Y.; Kliem, B.; Zhang, J.; Liu, Z.; Bi, Y.; Xiang, Y.; et al. Observing the Release of Twist by Magnetic Reconnection in a Solar Filament Eruption. *Nat. Commun.* **2016**, *7*, 11837. [[CrossRef](#)] [[PubMed](#)]
8. Berghmans, D.; Auchère, F.; Long, D.M.; Soubrié, E.; Mierla, M.; Zhukov, A.N.; Schühle, U.; Antolin, P.; Harra, L.; Parenti, S.; et al. Extreme-UV Quiet Sun Brightenings Observed by the Solar Orbiter/EUI. *Astron. Astrophys.* **2021**, *656*, L4. [[CrossRef](#)]
9. Cirtain, J.W.; Golub, L.; Winebarger, A.R.; De Pontieu, B.; Kobayashi, K.; Moore, R.L.; Walsh, R.W.; Korreck, K.E.; Weber, M.; McCauley, P.; et al. Energy Release in the Solar Corona from Spatially Resolved Magnetic Braids. *Nature* **2013**, *493*, 501–503. [[CrossRef](#)] [[PubMed](#)]
10. Regnier, S.; Alexander, C.E.; Walsh, R.W.; Winebarger, A.R.; Cirtain, J.; Golub, L.; Korreck, K.E.; Mitchell, N.; Platt, S.; Weber, M.; et al. Sparkling EUV Bright Dots Observed with Hi-C. *Astrophys. J.* **2014**, *784*, 134. [[CrossRef](#)]
11. Kuznetsov, V.D. Space Solar Research: Achievements and Prospects. *Phys.-Uspekhi* **2015**, *58*, 621. [[CrossRef](#)]
12. Barczynski, K.; Peter, H.; Savage, S.L. Miniature Loops in the Solar Corona. *Astron. Astrophys.* **2017**, *599*, A137. [[CrossRef](#)]
13. Williams, T.; Walsh, R.W.; Winebarger, A.R.; Brooks, D.H.; Cirtain, J.W.; Pontieu, B.D.; Golub, L.; Kobayashi, K.; McKenzie, D.E.; Morton, R.J.; et al. Is the High-Resolution Coronal Imager Resolving Coronal Strands? Results from AR 12712. *Astrophys. J.* **2020**, *892*, 134. [[CrossRef](#)]
14. Rahman, S.; Moon, Y.-J.; Park, E.; Siddique, A.; Cho, I.-H.; Lim, D. Super-Resolution of SDO/HMI Magnetograms Using Novel Deep Learning Methods. *Astrophys. J. Lett.* **2020**, *897*, L32. [[CrossRef](#)]
15. Yang, Q.; Chen, Z.; Tang, R.; Deng, X.; Wang, J. Image Super-Resolution Methods for FY-3E X-EUVI 195 Å Solar Images. *Astrophys. J. Suppl. Ser.* **2023**, *265*, 36. [[CrossRef](#)]
16. Bi, Y.; Yang, J.-Y.; Qin, Y.; Qiang, Z.-P.; Hong, J.-C.; Yang, B.; Xu, Z.; Liu, H.; Ji, K.-F. Morphological Evidence for Nanoflares Heating Warm Loops in the Solar Corona. *Astron. Astrophys.* **2023**, *679*, A9. [[CrossRef](#)]
17. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Loy, C.C.; Qiao, Y.; Tang, X. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018.
18. Lowe, D.G. Object Recognition from Local Scale-Invariant Features. In Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–27 September 1999; Volume 2, pp. 1150–1157.
19. Ariav, I.; Cohen, I. Fully Cross-Attention Transformer for Guided Depth Super-Resolution. *Sensors* **2023**, *23*, 2723. [[CrossRef](#)] [[PubMed](#)]

20. 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 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1874–1883.
21. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
22. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-Excitation Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019.
23. Barron, J.T. A General and Adaptive Robust Loss Function. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019.
24. Lai, W.-S.; Huang, J.-B.; Ahuja, N.; Yang, M.-H. Fast and Accurate Image Super-Resolution with Deep Laplacian Pyramid Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 2599–2613. [[CrossRef](#)] [[PubMed](#)]
25. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “Completely Blind” Image Quality Analyzer. *IEEE Signal Process. Lett.* **2013**, *20*, 209–212. [[CrossRef](#)]
26. 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.
27. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 295–307. [[CrossRef](#)] [[PubMed](#)]
28. Lim, B.; Son, S.; Kim, H.; Nah, S.; Lee, K.M. Enhanced Deep Residual Networks for Single Image Super-Resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Honolulu, HI, USA, 21–26 July 2017.
29. 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.
30. 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.
31. Li, J.; Fang, F.; Mei, K.; Zhang, G. Multi-Scale Residual Network for Image Super-Resolution. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 517–532.

**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.