



Improving the Spatial Resolution of Solar Images Based on an Improved Conditional Denoising Diffusion Probability Model

Wei Song^{1,2,3} , Wen Ma¹, Ying Ma¹, Xiaobing Zhao^{1,3}, and Ganghua Lin^{2,4}

¹ School of Information Engineering, Minzu University of China, Beijing 100081, People's Republic of China; songwei@muc.edu.cn

² Key Laboratory of Solar Activity, Chinese Academy of Sciences (KLSA, CAS), Beijing 100081, People's Republic of China

³ National Language Resource Monitoring and Research Center of Minority Languages, Minzu University of China, Beijing 100081, People's Republic of China

⁴ National Astronomical Observatories, Chinese Academy of Sciences (NAOC, CAS), Beijing 100081, People's Republic of China

Received 2022 May 27; revised 2022 October 7; accepted 2022 October 12; published 2022 November 22

Abstract

The quality of solar images plays an important role in the analysis of small events in solar physics. Therefore, the improvement of image resolution based on super-resolution (SR) reconstruction technology has aroused the interest of many researchers. In this paper, an improved conditional denoising diffusion probability model (ICDDPM) based on the Markov chain is proposed for the SR reconstruction of solar images. This method reconstructs high-resolution (HR) images from low-resolution images by learning a reverse process that adds noise to HR images. To verify the effectiveness of the method, images from the Goode Solar Telescope at the Big Bear Solar Observatory and the Helioseismic and Magnetic Imager (HMI) on the Solar Dynamics Observatory are used to train a network, and the spatial resolution of reconstructed images is 4 times that of the original HMI images. The experimental results show that the performance based on ICDDPM is better than the previous work in subject judgment and object evaluation indexes. The reconstructed images of this method have higher subjective vision quality and better consistency with the HMI images. And the structural similarity and rms index results are also higher than the compared method, demonstrating the success of the resolution improvement using ICDDPM.

Unified Astronomy Thesaurus concepts: Solar photosphere (1518); Neural networks (1933); Solar active regions (1974)

Supporting material: animation

1. Introduction

Solar activity is closely related to human production and life (Svalgaard 2013). Changes in solar activity affect the environment and climate on Earth (Feulner & Rahmstorf 2010; Engels & Van Geel 2012; Bamba et al. 2014). In addition, solar activity can also cause strong geomagnetic storms in the Earth's magnetosphere and ionosphere, which may interfere with the propagation of radio waves and the stability of navigation systems (Usoskin 2017). Predicting the level of solar activity can provide early warning for national security, aerospace, power equipment protection, etc. Therefore, the study of solar activity plays an important practical role.

The active regions of the Sun have a wide range of large-scale (such as sunspots and pores) and small-scale (such as Ellerman bombs) structures. Many small-scale events have the same characteristics as large-scale events. The study of small events and the comparison of the morphology of small events and large events are helpful for the analysis and prediction of solar activity (Nelson et al. 2013). A large number of fine structures used in the study of the solar active region are in a scale of 1'' (Chae et al. 2013; Wang & Liu 2019). Therefore, high-resolution (HR) solar images are beneficial for accurately measuring physical phenomena (Wang & Liu 2019). The continuous increase in the aperture of ground-based telescopes

has improved the angular resolution of the observed solar images to a certain extent. However, it also increases the cost of manufacturing and maintenance of the telescope. Moreover, due to the disturbance of atmospheric turbulence, the observed images of solar activity obtained by ground-based telescopes may still be fuzzy (Jia et al. 2019). In addition, a large amount of observed solar data has been accumulated. This series of historical data may be of great significance to the study of the solar cycle or solar activity (Ajabshirizadeh et al. 2011; Jonas et al. 2018). However, many events in active regions are difficult to obtain due to the limitation in image clarity. It has a certain impact on the study of historical data (Nelson et al. 2013). Therefore, it is necessary to use image postprocessing techniques for super-resolution (SR) reconstruction of low-resolution (LR) solar images.

The mapping relationship between images observed by different telescopes can be constructed to perform SR reconstruction of LR images, thereby providing clearer and richer data for the study of solar physics. The Goode Solar Telescope (GST) at the Big Bear Solar Observatory (BBSO) is a large ground-based telescope that can obtain solar images with a pixel size of 0''034, providing data for the study of small-scale structures (Goode & Cao 2012; Abramenko & Yurchyshyn 2020). However, GST has a limited field of view and cannot continuously observe the Sun for a long time. The Helioseismic and Magnetic Imager (HMI) on the launched Solar Dynamics Observatory (SDO) can acquire full-disk images of the Sun with a pixel size of 0''5 and provide uninterrupted observations of the Sun (Schou et al. 2012). But its spatial resolution is not sufficient to observe the small-scale structure of the Sun (Rahman et al. 2020). GST can provide HMI with more detailed information about the solar images

⁵ The second and third authors contributed equally to this work.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

during the same period. Therefore, a method based on deep learning is used to learn the mapping relationship between the HMI images and the GST images in this paper, thereby improving the spatial resolution of the HMI images.

The rest of the paper is organized as follows: Section 2 introduces three methods for SR reconstruction of solar images. The basic principles and network structure of our model are presented in Section 3. Section 4 shows the results of the SR reconstruction of our model, and quantitatively analyzes the results using the peak signal-to-noise ratio (peak S/N), structural similarity (SSIM), rms, power spectrum curves, scatter plot and gray histogram, etc., and also compares with Deng et al. (2021). We abbreviate the method in that paper as “SAGAN” later.

2. Related Work

The SR reconstruction of solar images refers to the restoration of the original clear images as accurately as possible from single or multiple LR images. Existing methods for SR reconstruction of solar images can be divided into methods based on interpolation, imaging models, and deep learning.

The commonly used methods based on interpolation include nearest neighbor interpolation and bicubic interpolation, etc. These methods have the advantages of low computational cost and high speed. However, the methods based on interpolation reconstruct the image based on the pixel values of the original image itself, so the reconstructed result may appear smooth or miss small-scale textures and other phenomena (Stengel et al. 2020; Xie et al. 2022). The methods based on the imaging model perform reconstruction tasks by obtaining the point-spread function (PSF), including speckle imaging, multiframe blind deconvolution (MFBD), phase diversity (PD), etc. The speckle imaging uses the statistical characteristics of atmospheric turbulence to reconstruct the phase and amplitude of the object (Wang et al. 2021). This method has been continuously developed since it was proposed in 1970, and it has also been used on solar telescopes such as GST and New Vacuum Solar Telescope (NVST). At present, a new HR reconstruction algorithm for solar images has also been proposed using its theory (Liu et al. 2022). Considering that PD and phase-diverse speckle (PDS) are special cases of MFBD, Löfdahl (2002) proposes a general formulation and uses a single algorithm to solve the MFBD problem. This method was successfully applied to the test data of PD, PDS, and MFBD. Multiobject multiframe blind deconvolution (MOMFBD) is a kind of MFBD. More information can be used to restore image details through MOMFBD with PD (Löfdahl et al. 2021). This kind of method can obtain good image quality by using the information from multiple images and different prior assumptions or prior values for PSF or the observed image (Ramos et al. 2018; Jia et al. 2019), and has been widely used in SR reconstruction of solar images (Ramos & Olsper 2021; Wang et al. 2021).

The methods based on deep learning reconstruct HR solar images by learning the mapping relationship between HR and LR images (Tian et al. 2019). These methods have been widely used in the SR reconstruction of solar images in recent years and have achieved good results. Baso & Ramos (2018) introduced deep learning earlier to the task of SR reconstruction of solar images based on two deep fully convolutional neural networks. Ramos et al. (2018) effectively combined deep learning and blind deconvolution processes, and proposed an end-to-end method for MFBD of solar images based on deep

convolutional neural networks. This method can reconstruct images quickly and has better real-time performance. Since solar images of the same wavelength have similar textures, Jia et al. (2019) used a cycle-consistent adversarial network to recover blurred images of the same stable physical process. Baso et al. (2019) recovered weak signals from images corrupted by complex noise through a neural network without prior knowledge of clean signals. This method is especially suitable for sparse wavelength sampling. Rahman et al. (2020) applied a residual attention model and a progressive generative adversarial network model to enhance the magnetograms of SDO/HMI. The magnetograms generated by the models are almost consistent with the corresponding target magnetograms. Jia et al. (2020) proposed a PSF-NET, which had a recurrent convolutional neural network structure. After training, PSF-NET is able to recover short-exposure images blurred by atmospheric turbulence to have the same contour. On the basis of this algorithm, Jia et al. (2021) further developed a deep neural network-based PSF modeling method. It trained the network with the collected system calibration data, and then estimated the PSF in any field of view from several discretely sampled star maps. Wang et al. (2021) proposed a channel-sharing spatiotemporal network, which used a complex spatial variant end-to-end blind restoration network to construct clear solar images. It is proved that this method cannot only remove the spatially variant blur included in the training, but also remove the unseen spatially variant blur, and has good generalization performance. Li et al. (2021) proposed a deblurring method based on coupling double-stage feature pyramid networks (FPN) with a single pipeline (DSFSP) to reconstruct HR solar speckle images. It can restore high-frequency details and remove artifacts. Deng et al. (2021) introduced a self-attention (SA) mechanism in conditional generative adversarial networks (CGAN) and constructed the HMI image and GST image pairs for training. This method can reconstruct HMI images with four times higher resolution.

The methods based on deep learning can learn a lot of information about HR images through a neural network and reconstruct high-quality images. This kind of method can also improve the quality of the image to a certain extent based on the data-driven idea (Wang et al. 2022). Therefore, the exploration of the methods based on deep learning has a good prospect. This study is inspired by denoising diffusion probabilistic models (DDPMs; Ho et al. 2020). DDPMs are generative models inspired by nonequilibrium thermodynamics that can generate high-quality images from noise and have a relatively complete theoretical derivation. DDPMs, normalizing flows (Rezende & Mohamed 2015; Baso et al. 2022), and variational autoencoder (Kingma & Welling 2013) are all likelihood-based generative models that use log-likelihood (or a suitable surrogate) as the training objective and can generate samples from the learned posterior. They have all been successfully applied to some tasks in the image processing (Song & Ermon 2019; Whang et al. 2021). DDPMs have been proven to outperform state-of-the-art models of the same period in sample quality in image generation (Dhariwal & Nichol 2021). Compared with the state-of-the-art GAN-based methods of the same period, the outputs of DDPMs have better consistency with LR inputs in image SR reconstruction (Saharia et al. 2022). And DDPMs have also achieved good results in other image fields (Choi et al. 2021; Whang et al. 2021; Lugmayr et al. 2022). This probabilistic model adds Gaussian noise to images in multiple

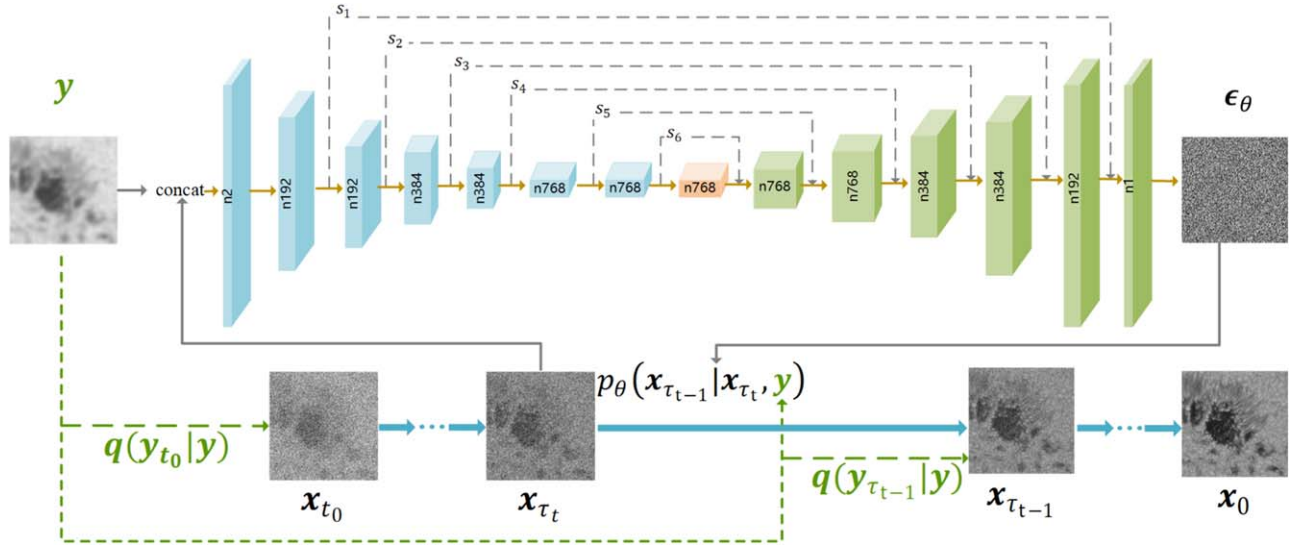


Figure 1. The overall framework of our ICDDPM. The number after n is the number of the channel. y , x_0 , and ϵ_θ represent the HMI image, GST image, and prediction noise, respectively.

steps to gradually convert a data distribution into a Gaussian distribution. The model then learns the inversion of this process and iteratively estimates the reversal results for multiple steps, which is simpler than estimating the data distribution directly (Sohl-Dickstein et al. 2015). It is also allowed to refine and correct the results of a single step after each estimation. This controls the reconstruction process so that the results are toward the desired target (Choi et al. 2021). In terms of model optimization, this model is easy to extend (Dhariwal & Nichol 2021) and modify (Song et al. 2020). Considering the theoretical basis, good performance in many fields, controllable reconstruction process, and large optimization space of DDPMs, this paper proposes an improved conditional denoising diffusion probability model (ICDDPM) based on the principle of conditional DDPMs. And this model is applied to the SR reconstruction of solar images.

The contributions of our paper are as follows:

1. DDPMs are applied to the SR reconstruction of solar images. Based on the conditional DDPMs, the initialization and the middle of the reconstruction process are improved to enhance the consistency between the reconstruction results and the input images;
2. A new auxiliary term based on contrastive regularization is added to the loss function to improve the reconstruction ability of the model, so that the reconstruction results have a certain improvement in the indexes.

3. An Improved Conditional Denoising Diffusion Probability Model

3.1. Basic Principles

An ICDDPM transforms the original distribution of the HR solar image x_0 into a Gaussian distribution by gradually adding Gaussian noise to x_0 . This process is a forward process. Under the action of the LR solar image y , the reversal of the forward process is learned so that the noisy image is denoised in multiple steps, and obtains the SR reconstructed image. This process is a reverse process (Figure 1). This section specifically

describes the forward and reverse processes, and introduces the loss function used to train the model.

3.1.1. Forward Process

Given the distribution $x_0 \sim q(x_0)$ of the HR solar image. The Gaussian noise ϵ is gradually added to the solar image x_0 by T steps, so that the latents $x_1, \dots, x_T$ are obtained sequentially. x_T is approximately an isotropic Gaussian distribution. The forward process is fixed on a Markov chain with T transitions so that the original data distribution is transformed into a known simple distribution (Ho et al. 2020; Nichol & Dhariwal 2021):

$$q(x_{1:T}|x_0) = \prod_{t=1}^T q(x_t|x_{t-1}),$$

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t \mathbf{I}), \quad (1)$$

where $\beta_1, \dots, \beta_T$ is an increasing sequence in the range of (0,1), which is used to control the noise addition rate of the forward process. Ho et al. (2020) set β_t to a small value, that is, adding noise in small steps. This setting makes when the forward $q(x_t|x_{t-1})$ is a Gaussian, and the reverse $q(x_{t-1}|x_t)$ is also a Gaussian (Xiao et al. 2021). The property of the forward process allows obtaining a noisy solar image x_t at any t given x_0 in a closed form. This property avoids iteratively solving x_t from x_0 :

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1 - \bar{\alpha}_t)\mathbf{I}), \quad (2)$$

where $\alpha_t = 1 - \beta_t$, $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$. By Equation (2), the relationship between the noisy solar image x_t , the HR solar image x_0 , and the noise ϵ can be obtained:

$$x_t(x_0, \epsilon) = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{(1 - \bar{\alpha}_t)}\epsilon, \quad (3)$$

where $\epsilon \sim \mathcal{N}(0, \mathbf{I})$. According to Equations (1) and (2), using Bayes' theorem, the posterior distribution of x_{t-1} can be obtained when x_t and x_0 are given in the forward process

(Ho et al. 2020; Song et al. 2020; Nichol & Dhariwal 2021):

$$\begin{aligned} q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) &= \mathcal{N}(\mathbf{x}_{t-1}; \tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0), \tilde{\boldsymbol{\beta}}_t \mathbf{I}), \\ \tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0) &= \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1 - \bar{\alpha}_t} \mathbf{x}_0 + \frac{\sqrt{\bar{\alpha}_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} \mathbf{x}_t, \\ \tilde{\boldsymbol{\beta}}_t &= \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} \beta_t. \end{aligned} \quad (4)$$

The reverse process fits this posterior distribution to complete the denoising operation. The principle of the reverse process will be explained later.

3.1.2. Reverse Process

The model learns the reversal of the forward process, and HR solar images can be reconstructed through iterations of the reverse process. The reverse process is also fixed on a Markov chain. Many DDPMs start with $p(\mathbf{x}_T) \sim \mathcal{N}(\mathbf{x}_T; \mathbf{0}, \mathbf{I})$ (Saharia et al. 2022; Li et al. 2022). However, this study draws on the idea of Stochastic Differential Editing (SDEdit; Meng et al. 2021). A hyperparameter t_0 ($t_0 < T$) is selected in the reverse process, and the iteration starts from t_0 (Figure 1). This has two advantages. First, it reduces the number of iterations from T to t_0 , which saves time in reconstructing solar images. Second, the result of Gaussian perturbation of the LR image $\mathbf{y}$ at time t_0 is used as the initialization of the reconstruction. It adds more information about the solar image than starting from a Gaussian distribution. The improved initialization is expressed as:

$$\mathbf{x}_{t_0} = \sqrt{\bar{\alpha}_{t_0}} \mathbf{y} + \sqrt{1 - \bar{\alpha}_{t_0}} \boldsymbol{\epsilon}. \quad (5)$$

Under the condition $\mathbf{y}$, the reverse process learns $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{y})$ to fit $q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)$, thereby guiding the reconstructed image to be close to the HR image corresponding to the LR image. Finally, $p_\theta(\mathbf{x}_0)$ obtained by the reverse iteration is used to fit the real distribution $q(\mathbf{x}_0)$ of the solar image:

$$p_\theta(\mathbf{x}_{0:t_0}|\mathbf{y}) = p(\mathbf{x}_{t_0}) \prod_{t=1}^{t_0} p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{y}), \quad (6)$$

where $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{y})$ can be expressed as a Gaussian transition with a learned mean and a fixed variance $\Sigma_\theta(\mathbf{x}_t, \mathbf{y}, t) = \sigma_t^2 \mathbf{I}$:

$$\begin{aligned} p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{y}) \\ = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, \mathbf{y}, t), \Sigma_\theta(\mathbf{x}_t, \mathbf{y}, t)). \end{aligned} \quad (7)$$

According to Equations (3) and (4), $\boldsymbol{\mu}_\theta$ can be decomposed into a linear combination of $\mathbf{x}_t$ and a prediction noise $\boldsymbol{\epsilon}_\theta$. Denoising diffusion implicit models (DDIMs; Song et al. 2020) remove the random item by setting σ_t to 0. This operation can increase the certainty of the reverse process. Since the task of SR reconstruction of solar images should reduce the uncertainty of the reconstruction process, this paper also adopts the same setting. This enables the HR solar image to be reconstructed from the latent in fixed processes. That is, under the same initialization, it is avoided that the detailed features of the results obtained by multiple reconstructions are quite different. The reconstruction process of the solar image is expressed as:

$$\begin{aligned} \mathbf{x}_{t-1} &= \boldsymbol{\mu}_\theta(\mathbf{x}_t, \mathbf{y}, t) \\ &= \frac{1}{\sqrt{\bar{\alpha}_t}} (\mathbf{x}_t - \frac{\beta_t}{\sqrt{1 - \bar{\alpha}_t}} \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{y}, t)). \end{aligned} \quad (8)$$

In addition, accelerated sampling methods are used. It can further accelerate the reconstruction process without compromising the quality of the reconstruction results (Song et al. 2020). That is, in the process of $1 \sim t_0$, a subset $\mathbf{x}_{\tau_1}, \dots, \mathbf{x}_{\tau_S}, \dots, \mathbf{x}_{\tau_S}$ of length S ($S < t_0$) is extracted and the corresponding noise schedule is selected, where τ is an increasing subsequence of $[1, \dots, t_0]$. Thus, after S steps of iteration from $\mathbf{x}_{\tau_i}$ to $\mathbf{x}_{\tau_{i+1}}$, the reconstructed image of the LR image $\mathbf{y}$ can be obtained.

In order to make the reconstruction result clear and faithful to the LR input, it is necessary to pay attention to the consistency between the input and output (Saharia et al. 2022). Therefore, the intermediate steps of the reverse process of conditional DDPMs are improved in this paper. This paper introduces the idea of Iterative Latent Variable Refinement (ILVR; Choi et al. 2021). The generated $\mathbf{x}_{\tau_{i+1}}$ after each iteration is refined by $\mathbf{y}$ and used as the input of the next iteration. This ensures that the downsampled image $\phi_N(\mathbf{x}_0)$ of the reconstructed image $\mathbf{x}_0$ is equal to $\phi_N(\mathbf{y})$. That is, the target reconstructed image has the same low-frequency information as the LR image. $\phi_N(\cdot)$ represents linear low-pass filtering, which is a sequence of downsampling and upsampling by a factor of N .

According to the linear properties of ϕ_N and Equation (3), $\phi_N(\mathbf{x}_{\tau_{i+1}}) = \phi_N(\mathbf{y}_{\tau_{i+1}})$, $\mathbf{y}_{\tau_{i+1}} = \sqrt{\bar{\alpha}_{\tau_{i+1}}} \mathbf{y} + \sqrt{1 - \bar{\alpha}_{\tau_{i+1}}} \boldsymbol{\epsilon}$ can be obtained. Therefore, the distribution of $\mathbf{x}_{\tau_{i+1}}$ can be refined by matching $\phi_N(\mathbf{x}_{\tau_{i+1}})$ of $\mathbf{x}_{\tau_{i+1}}$ with $\phi_N(\mathbf{y}_{\tau_{i+1}})$ of $\mathbf{y}_{\tau_{i+1}}$:

$$\mathbf{x}_{\tau_{i+1}} = \phi_N(\mathbf{y}_{\tau_{i+1}}) + \mathbf{x}_{\tau_{i+1}} - \phi_N(\mathbf{x}_{\tau_{i+1}}). \quad (9)$$

In summary, the LR solar image $\mathbf{y}$ has three purposes: the first is to use the $\mathbf{y}_{t_0}$ obtained from $\mathbf{y}$ as the initialization $\mathbf{x}_{t_0}$ of the reverse process to increase the useful information of initialization; the second is that it acts as a condition to guide the reverse process; the third is to refine each $\mathbf{x}_{\tau_{i+1}}$, thereby limiting the reconstruction result to be faithful to the corresponding LR image. Therefore, with reasonable initialization, the guidance of the LR images, and refinement of each reconstruction step, the reconstruction result with enhanced consistency with the input $\mathbf{y}$ can be obtained.

3.1.3. Loss Function

In order to enable the network to better complete the mapping between the LR and HR solar images, the model needs to be controlled by the loss function. During model training, $\boldsymbol{\mu}_\theta(\mathbf{x}_t, \tau_t)$ in $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{y})$ needs to fit the true reverse distribution $\tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0)$ in $q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)$. So the loss function is expressed as:

$$L = \mathbb{E}_q[\|\tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0) - \boldsymbol{\mu}_\theta(\mathbf{x}_t, \tau_t)\|^2]. \quad (10)$$

According to the representation of the mean in Equation (8), the loss function can be simplified (Ho et al. 2020):

$$\begin{aligned} L &= \mathbb{E}_{\tau_t, \mathbf{x}_0, \mathbf{y}, \boldsymbol{\epsilon}}[\|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_{\tau_t}} \mathbf{x}_0 \\ &\quad + \sqrt{1 - \bar{\alpha}_{\tau_t}} \boldsymbol{\epsilon}, \mathbf{y}, \tau_t)\|^2]. \end{aligned} \quad (11)$$

In order to enhance the ability of the model to predict HR images $\mathbf{x}_0$, this paper discusses a new loss function term on the basis of retaining the existing loss function. Wu et al. (2021) proves that the contrastive regularization can be well applied to the problem of improving image sharpness. This idea is used in this task of SR reconstruction of solar images. It can be explained as follows: combining the $\boldsymbol{\epsilon}_\theta$ and the known $\mathbf{x}_{\tau_t}$ at each τ_t time, the prediction $\tilde{\mathbf{x}}_0$ of $\mathbf{x}_0$ at this time can be obtained

according to Equation (3). The loss calculation involving only the predicted $\tilde{x}_0$ and x_0 cannot use the information of the LR image y . Thus, the loss calculation that makes $\tilde{x}_0$ close to the details of x_0 and far away from the details of y is considered. It can make full use of the known information of HR and LR images. The extracted features of each image are used to participate in the loss calculation:

$$L_{\text{VGG}} = L + \lambda \sum_{i=1}^n \omega_i \cdot \frac{D(G_i(x_0), G_i(\tilde{x}_0))}{D(G_i(y), G_i(\tilde{x}_0))}, \quad (12)$$

where, λ is a hyperparameter to prevent the added auxiliary loss term from overwhelming L , and is set to 0.1. The value of n is 5, which represents five layers of the pretrained VGG-19 network. The hidden features of the images extracted by these five layers are used. ω_i is a coefficient that controls the weight of features of each layer. $D(I, J)$ represents the L1 distance between I and J . G_i represents the i th extracted feature. This paper uses the same hyperparameters and pretrained VGG-19 as Wu et al. (2021).

L_{VGG} helps improve the indexes of some results. In order to further improve the results on multiple evaluation indexes, the images are directly used for loss calculation. The loss function is further improved as:

$$L_{\text{new}} = L + \lambda \frac{D(x_0, \tilde{x}_0)}{D(y, \tilde{x}_0)}. \quad (13)$$

3.2. Network Structure

The network structure of ICDDPM refers to the design of Nichol & Dhariwal (2021), which is a U-Net backbone used to predict noise at different times. During training, the LR image y and its corresponding noisy HR image x_τ are concatenated as input, and the range of pixel values is normalized to $[-1, 1]$. To facilitate concatenation of y and x_τ based on channel, the LR image is subjected to 4 times bilinear interpolation. The predicted value ϵ_θ of the noise added to the HR image x_0 is the output (Figure 1). During testing, the LR image and its Gaussian perturbation results are concatenated and input into U-Net. By gradually removing the S prediction noises output by the network, the reconstructed HR image is finally obtained.

The network consists of three parts: the shrinking path, the middle part, and the expanding path. The shrinking path extracts the features of the input image, with a total of six layers:

$$M = S_6(S_5(S_4(S_3(S_2(S_1(x)))))), \quad (14)$$

where x , S_i and M represent the input of the network, the operation of the i th layer on the shrinking path and the input of the middle part, respectively. Each layer contains two residual blocks. Each residual block contains two operations consisting of a group normalization, an activation function and a filter of size 3 and stride 1. A transformer sinusoidal position embedding (Vaswani et al. 2017) is added in the middle of the two operations to make the residual block determine the time t . Residual blocks are used to mitigate network degradation (He et al. 2016). Group normalization can reduce the impact of batch size (Wu & He 2018). Except for S_6 , the other layers perform a downsampling operation. When the feature map size is 16×16 and 8×8 , the shrinking path S_5 and S_6 add the SA mechanism with four attention heads. The

multihead attention is beneficial for the model to pay attention to the information of different representation subspaces at different positions (Vaswani et al. 2017). The number of channels gradually increases according to the specified channel multiplier.

The middle part consists of two residual blocks and a SA module, whose output is the input of the expanding path. The expanding path is similar to the shrinking path, which also contains SA mechanisms. Dhariwal & Nichol (2021) found in experiments that increasing the number of the SA module and the SA head can improve the performance of the network. However, this network does not adopt a larger architecture due to the limitations of the experimental environment. The output of the upper layer and the feature map of the corresponding size of the shrinking path are skipped as the input of the lower layer. Each layer contains three residual blocks to receive feature maps of corresponding scales on the shrinking path. This maximizes the use of information on the shrinking path. The upsampling operation is performed in the first five layers, and the number of channels is gradually reduced. As shown in the following equation:

$$O_k = E_k(O_{k-1} + S_{N-k+1}), \quad (15)$$

where O_k , E_k and S_{N-k+1} represent the output result of the k th layer, the operation of the k th layer of the expanding path and the output of the corresponding layer of the shrinking path, respectively. N is set to 6, which means that the expanding path has six layers.

This network is implemented on Pytorch. Existing models based on DDPMs use more GPUs, making the batch size up to 256 (Nichol & Dhariwal 2021; Saharia et al. 2022). Our lab is limited by hardware and uses a V100, so the batch size is set to four. To match this batch size, the learning rate is chosen to be a small 1×10^{-5} . Although this limits the performance of the algorithm to a certain extent, it still achieves relatively good results. T is set to 4000, and this larger T is chosen to obtain a smaller β . $\beta_1, \dots, \beta_T$ are set as a linear increasing sequence from 2.5×10^{-5} to 5×10^{-3} (Nichol & Dhariwal 2021). When reconstructing the image, the optimal parameters are selected through experiments. The hyperparameter γ is set to 2.3 to perform brightness preprocessing on the test set. The hyperparameter t_0 is set to 3000. When t_0 is smaller or larger, the results are usually worse on the indexes. The number of reconstruction step S affects the reconstruction time. The experimental results show that a larger S does not significantly improve the results of each index while spending more reconstruction time. When S is set to smaller values (such as 25, 15, 10, and 5), the peak S/N and SSIM indexes of the results will get better with the decrease of S , but the rms index will get worse. Therefore, S is set to 25. In the analysis part of the experimental results, S is 25 as an example to show, unless otherwise specified. The specific details of the network are shown in Table 1.

4. Result Analysis

4.1. Data Set

The same scale-invariant feature transform (SIFT) method (Lowe 2004) for image pairing as SAGAN is used. The SIFT method is completed through three main steps of feature point identification, feature point matching, and registration parameter determination. The GST image and the HMI image are

Table 1
Structure of the Neural Network

Layer	Channel Multiplier	Self-attention	Number of Residual Blocks	Upsampling/Downsampling	Input (H × W × C)	Output (H × W × C)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
S1	1	0	2	down	256 × 256 × 2	128 × 128 × 192
S2	1	0	2	down	128 × 128 × 192	64 × 64 × 192
S3	2	0	2	down	64 × 64 × 192	32 × 32 × 384
S4	2	0	2	down	32 × 32 × 384	16 × 16 × 384
S5	4	2	2	down	16 × 16 × 384	8 × 8 × 768
S6	4	2	2	no	8 × 8 × 768	8 × 8 × 768
Middle	...	1	2	no	8 × 8 × 768	8 × 8 × 768
E1	4	3	3	up	8 × 8 × 768	16 × 16 × 768
E2	4	3	3	up	16 × 16 × 768	32 × 32 × 768
E3	2	0	3	up	32 × 32 × 768	64 × 64 × 384
E4	2	0	3	up	64 × 64 × 384	128 × 128 × 384
E5	1	0	3	up	128 × 128 × 384	256 × 256 × 192
E6	1	0	3	no	256 × 256 × 192	256 × 256 × 1

Note. The number of each module used in the different layers of the network structure and the dimensions of the input and output are described in detail in the table. The third column represents the number of times the SA mechanism with four attention heads. In the fifth column, “down” represents downsampling, “up” represents upsampling, and “no” represents no upsampling or downsampling.

taken as HR and LR image pairs, and SIFT matches the image pairs by the feature points of the HR and LR images. In order to avoid the large difference in resolution between the GST image and the HMI image from affecting the registration, the GST image needs to be preprocessed. The original GST image is Gaussian blurred and downsampled to $0''.126$. This spatial resolution is about four times that of the HMI image, and image patches of size 256×256 pixels are cropped from their centers.

The original data of ICDDPM and SAGAN are consistent with verifying the effectiveness of the algorithm. A total of 1597 GST and HMI image pairs are used, of which 1497 are used as a training set and the remaining 100 are used as a test set.

4.2. Evaluation Index

In order to evaluate the results of our model objectively, peak S/N, SSIM, and rms are introduced in this paper. Peak S/N is based on the error of the corresponding pixels of the HR image and the SR image. The higher the value of peak S/N, the higher the similarity between the two images. Peak S/N is a relatively broad objective evaluation index for images.

$$\text{Peak S/N} = 10 \log_{10} \left(\frac{M^2}{\frac{1}{t} \sum_{i=1}^t (I_{\text{HR}}(i) - I_{\text{SR}}(i))^2} \right), \quad (16)$$

where t , I_{HR} , and I_{SR} represent the total number of pixels of the image, the pixel value of the HR image, and the pixel value of the SR image, respectively. M is the maximum pixel value of the image, and the value is 255 for the sampling point represented by 8 bits.

Since peak S/N does not consider the structural information of the image, the evaluation index SSIM is also used. SSIM measures the similarity of images from three aspects:

brightness σ , contrast μ , and structure. It is defined as:

$$\text{SSIM} = \frac{(2\mu_{\text{HR}}\mu_{\text{SR}} + c_1)(\sigma_{\text{HRSR}} + c_2)}{(\mu_{\text{HR}}^2 + \mu_{\text{SR}}^2 + c_1)(\sigma_{\text{HR}}^2 + \sigma_{\text{SR}}^2 + c_2)}, \quad (17)$$

where $c_1 = 6.5025$ and $c_2 = 58.5225$. Since the peak S/N and SSIM indexes are calculated by the skimage library (Van der Walt et al. 2014) in the experiment, the two constants here are set directly by the code in the library without manual setting. Brightness and contrast are obtained from the covariance and mean of the images, respectively.

The rms is the most commonly used method in solar image quality evaluation (Popowicz et al. 2017; Denker et al. 2018). It is related to the contrast of the image, the higher the value, the higher the image quality. It is defined as:

$$\text{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N (I - \bar{I})^2} / \bar{I}, \quad (18)$$

where N , I , and $\bar{I}$ are the total number of pixels of the image, the pixel values of the image, and the average of image pixel values, respectively.

4.3. Comparative Experiment

To show the complete reconstruction process of a solar image, Figure 2 shows the intermediate output results of the reconstruction in turn. The examples shown below are the reconstruction results of models trained with L_{new} as the loss function. The contour of the sunspot has been clearly seen in the output of the fifth iteration. And the penumbra region has appeared in the output of the sixth iteration (Figure 2). The 100 images in the test set are SR reconstructed. Under the same hardware conditions, the reconstructed process in the test part based on ICDDPM consumes about 4 minutes, which is four times slower than the reconstruction by SAGAN. The reconstruction results for five HMI images are shown in Figure 3, with a field of view of about $30'' \times 30''$. The first two rows are examples of granulation in solar inactive regions and a

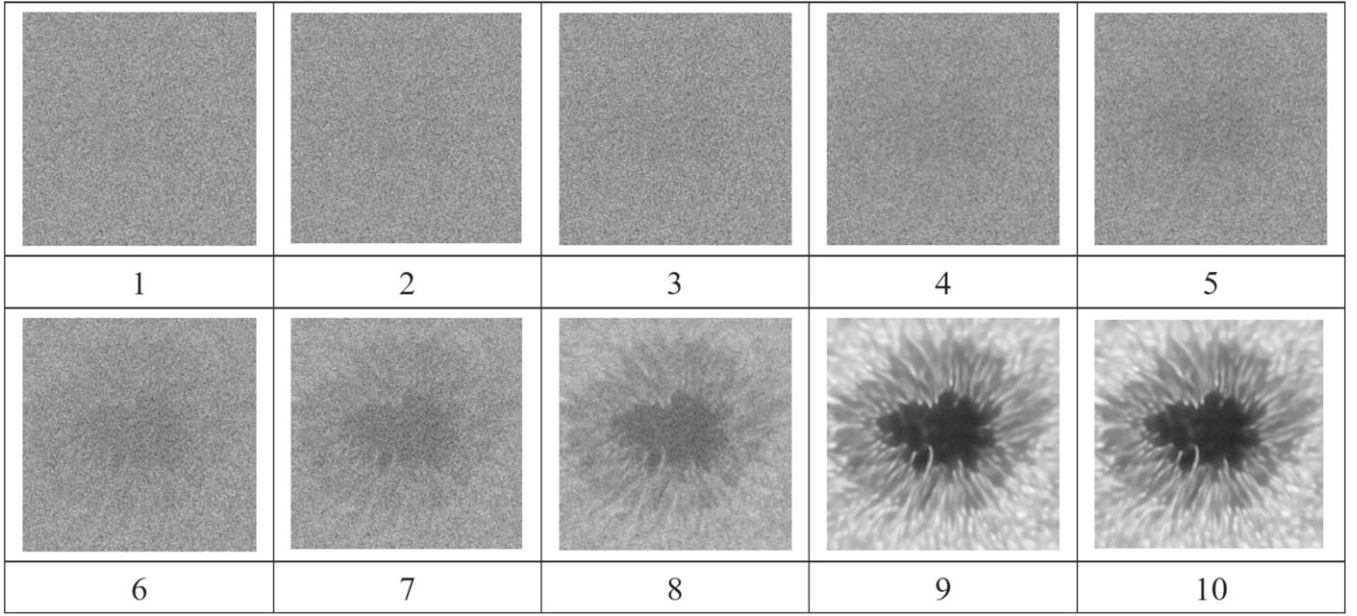


Figure 2. The reconstruction process of a solar image. The numbers below each image represent the number of iterations. The number of reconstruction step S is 10.

Table 2
The Average Value of Different Indexes of the Experimental Results

Evaluation Index (1)	HMI Image (2)	SAGAN SR Image (3)	Our SR Image (L) (4)	Our SR Image (L_{VGG}) (5)	Our SR Image (L_{new}) (6)
Peak S/N/dB	12.41220	23.01999	20.26631	19.61052	20.32623
SSIM	0.39692	0.47109	0.52120	0.50789	0.54558
rms	0.10842	0.22911	0.23284	0.26369	0.23294

Note. This table shows the average value of peak S/N and SSIM of HMI images, SAGAN SR images, our SR images with GST images in the test set, and the average of rms for various images.

small pore, and the last three rows are examples of solar active regions containing sunspots. From the perspective of subjective evaluation, the SR images of ICDDPM reconstruct a large number of details on the basis of preserving the HMI image structure. This result achieves our desired goal. As shown in the red boxes in Figure 3, our model reconstructs wider granule borders and penumbral filaments compared with SAGAN. But most of them are closer in location and orientation to the details in GST images. Although SAGAN reconstructs finer structures, it also introduces more artifacts. This may be because the two models use different loss functions. The adversarial loss in SAGAN emphasizes more on perceptual quality and can reconstruct more details. But it may also lead to the introduction of artifacts. The details on the image are irregular and human perception is not very sensitive to this (Liang et al. 2022b). Therefore, the details and artifacts reconstructed by SAGAN are difficult to distinguish. This may affect the application of reconstructed images in scientific fields. ICDDPM mainly employs a mean squared error loss. This loss emphasizes fidelity (Liang et al. 2022b) and thus may lead to wider granule borders and penumbral filaments being reconstructed. In addition, our model focuses on the consistency of the reconstructed image with the HMI image, while SAGAN focuses on the similarity of the reconstructed image with the GST image. Therefore, the results of ICDDPM are slightly closer to the HMI images in terms of brightness and contrast. Figure 4 shows the reconstructed details of the small-

scale structure of the solar image. It can be seen that our model is able to reconstruct structures smaller than $1''$ and obtain more natural details.

An animation of 833 images of the Sun is created in a continuous observation period (Figure 5). The animation shows a typical dynamic of a sunspot. From a subjective angle, both SAGAN and ICDDPM reconstruct many detailed structures. However, the edges of the reconstructed image of SAGAN have slightly obvious artifacts. And it has unnatural penumbral filaments in the lower left area of the penumbra. The details of the reconstructed image of ICDDPM are more natural.

In order to test the effect of different loss functions, this paper conducts comparative experiments on L , L_{VGG} , and L_{new} . The average values of peak S/N, SSIM, and rms on the test set are shown in Table 2. The results show that, compared with the values of the indexes of the HMI images, the results obtained by using different loss functions are improved to a certain extent. Comparing the results of the three indexes, the model with L_{new} as the loss function has better results. This may be because the added loss term constrains the upper and lower bounds of the result. This constraint enables the reconstruction results to contain details close to the GST images while staying away from the fuzzy information in the HMI images. Although L_{VGG} is also constrained by upper and lower bounds, it may be that the features extracted from the solar image involved in the loss calculation are inaccurate. This limits the improvement of the indexes.

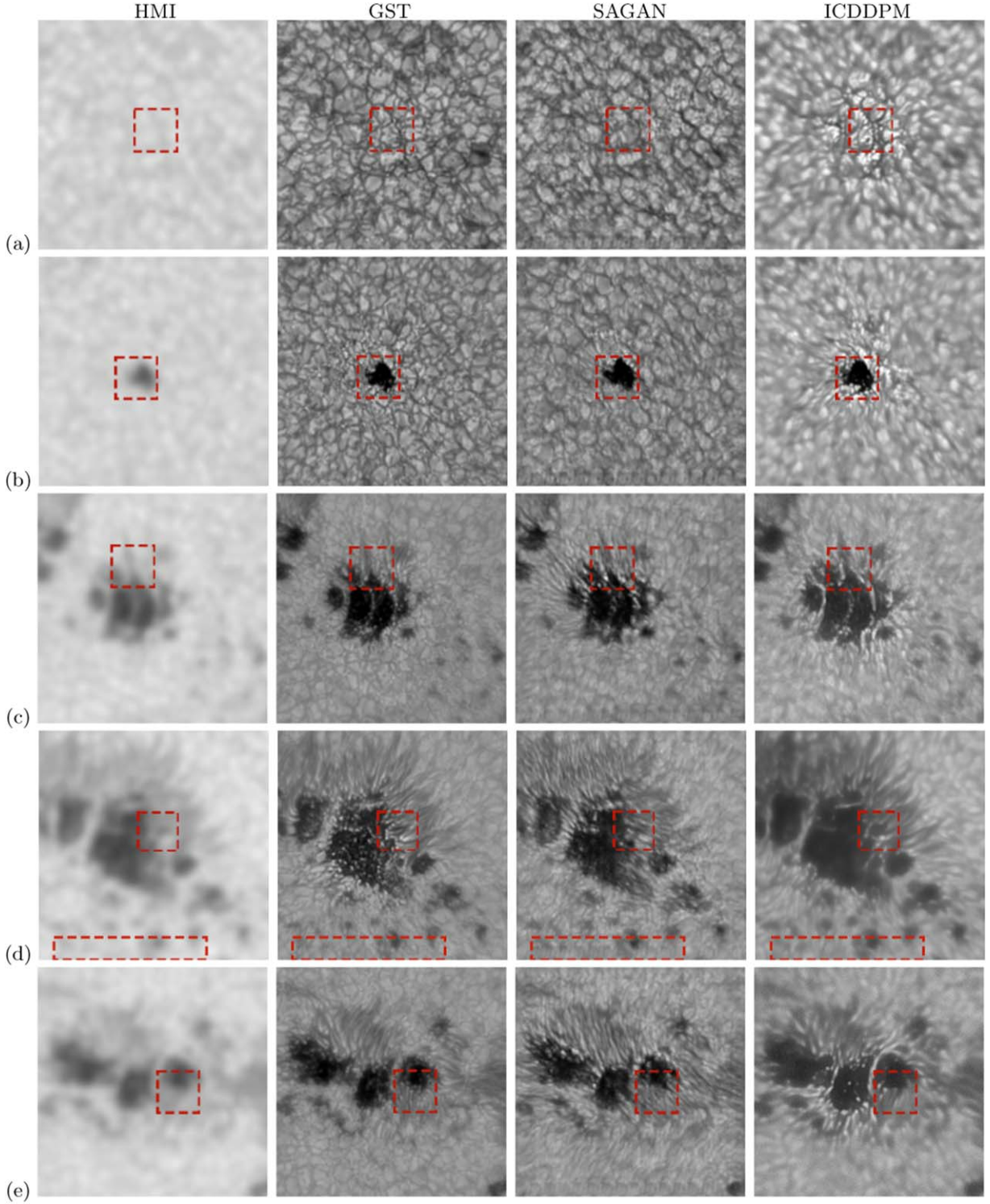


Figure 3. Five reconstruction results. The images from left to right correspond to the HMI image, the GST image, the SAGAN SR image, and our SR images, respectively. The field of view is about $30'' \times 30''$. Red dotted boxes are used to facilitate the comparison of images.

Figure 6 shows the comparison of the evaluation indexes for each reconstructed image of ICDDPM and SAGAN under the same test set. Figure 6(a) shows the pixel-level similarity between the reconstructed image and the GST image. The larger the value of peak S/N, the more similar the reconstructed image is to the GST image. SAGAN outperforms ICDDPM in similarity to GST images (Figure 6(a)). In brightness, our reconstruction results are closer to the HMI images. And the

reconstructed images of SAGAN are closer to the GST images. This may have affected the peak S/N results. Therefore further evaluations of the results are introduced. The SSIM index not only considers the influence of the brightness, but also the structure and contrast of the image. The reconstruction results of ICDDPM are generally slightly better than those of SAGAN in the SSIM index (Figure 6(b)). Compared with SAGAN, our results increase the average SSIM index and rms index on the

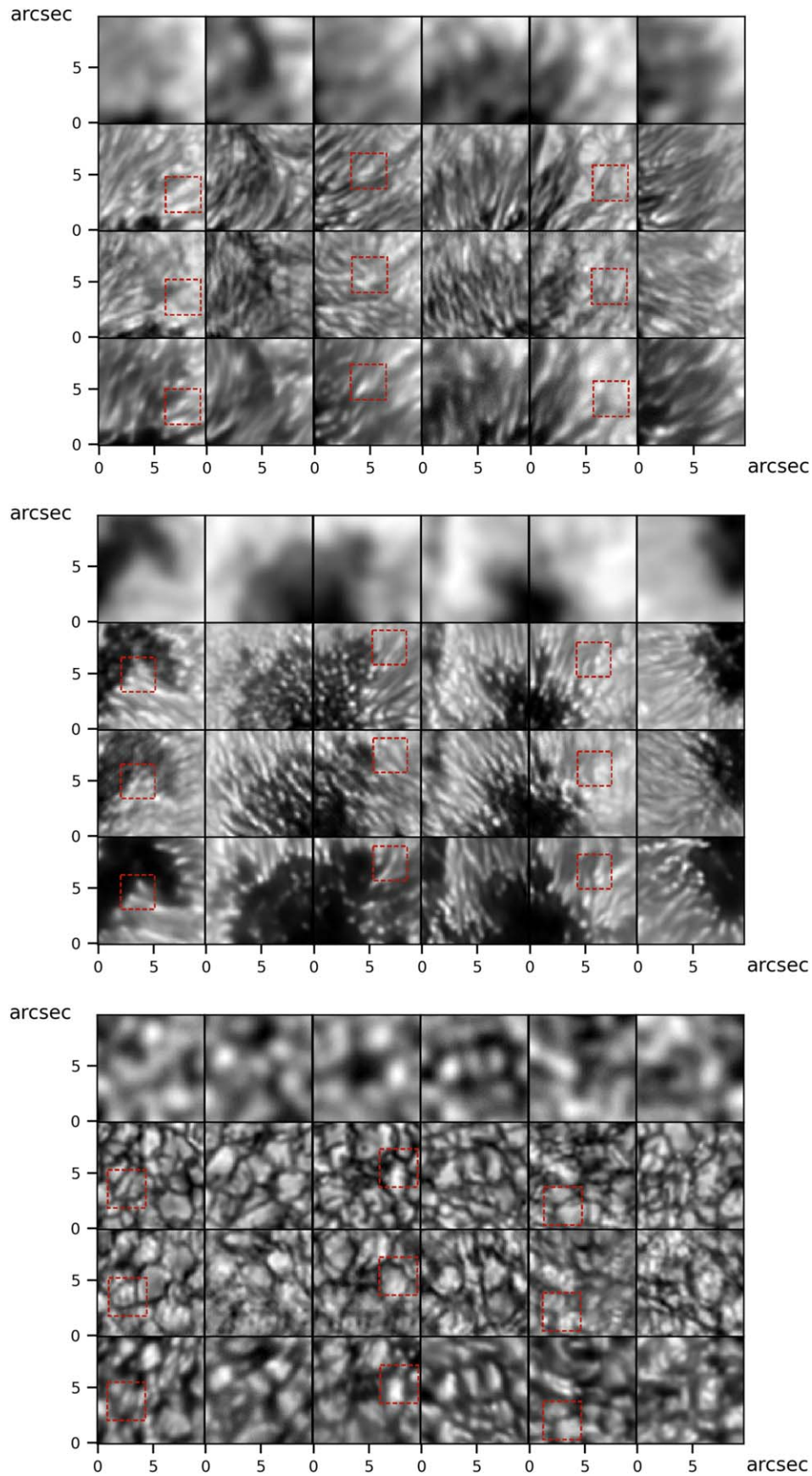


Figure 4. Shows the reconstructed small-scale structures of the solar penumbra (top), umbra (middle), and granulation (bottom) in turn. The first row is the HMI image, the second row is the GST image, the third row is the SR image of SAGAN, and the fourth row is our SR image. Red dotted boxes are used to facilitate the comparison of images.

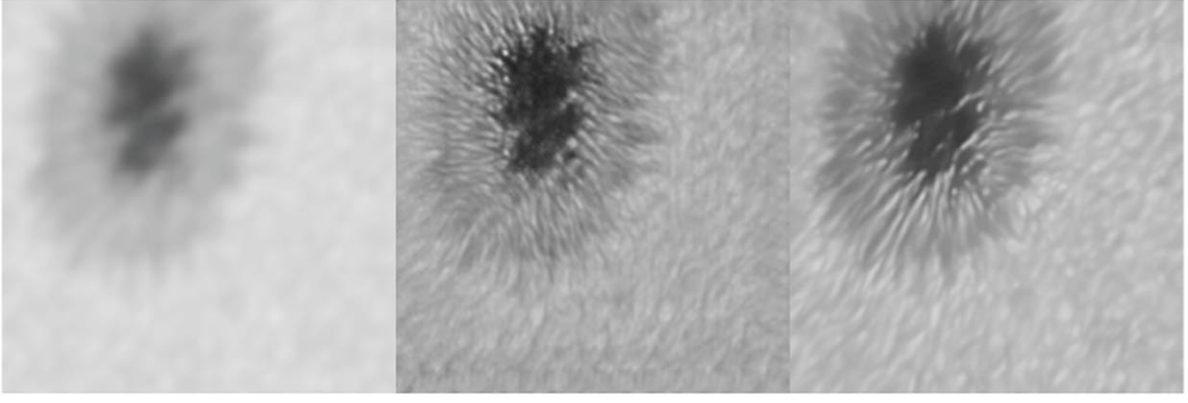


Figure 5. An animation of SR reconstruction results from 2017 September 1 to 9, comparing the HMI image (left), SAGAN SR image (middle), and our SR image (right). The animation runs for a total of 34 s, and the reconstruction shows a dynamic progression of a sunspot.

(An animation of this figure is available.)

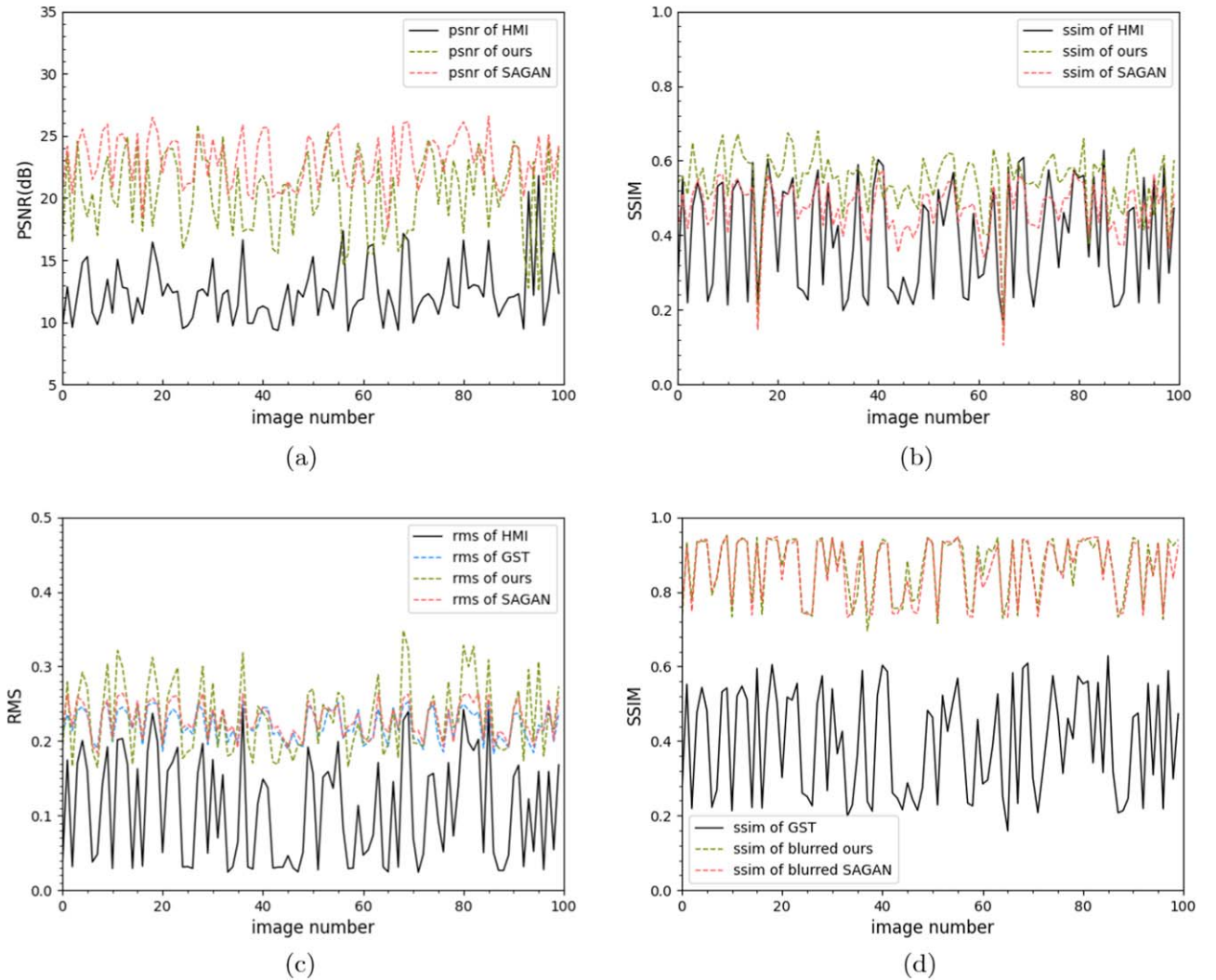


Figure 6. Peak S/N, SSIM, and rms indexes of test results. The black lines in (a) and (b) represent the indexes of the HMI image and GST image, while the green and red lines represent the indexes of our SR images, SR images of SAGAN and GST images, respectively. The black, blue, green, and red lines in (c) represent the indexes of the HMI image, GST image, our SR images, and SR images of SAGAN, respectively. (d) represents the SSIM index of the GST images, and the index of our SR images and the SR images of SAGAN after blurring.

test set by 0.074 and 0.004, respectively (Table 2). This may represent that the reconstructed images of ICDDPM have certain advantages in structure and contrast. In addition, on the

fluctuation of the rms curve, the SR image of SAGAN is similar to the GST image, while our SR image is similar to the HMI image (Figure 6(c)). This may indicate that the SR images

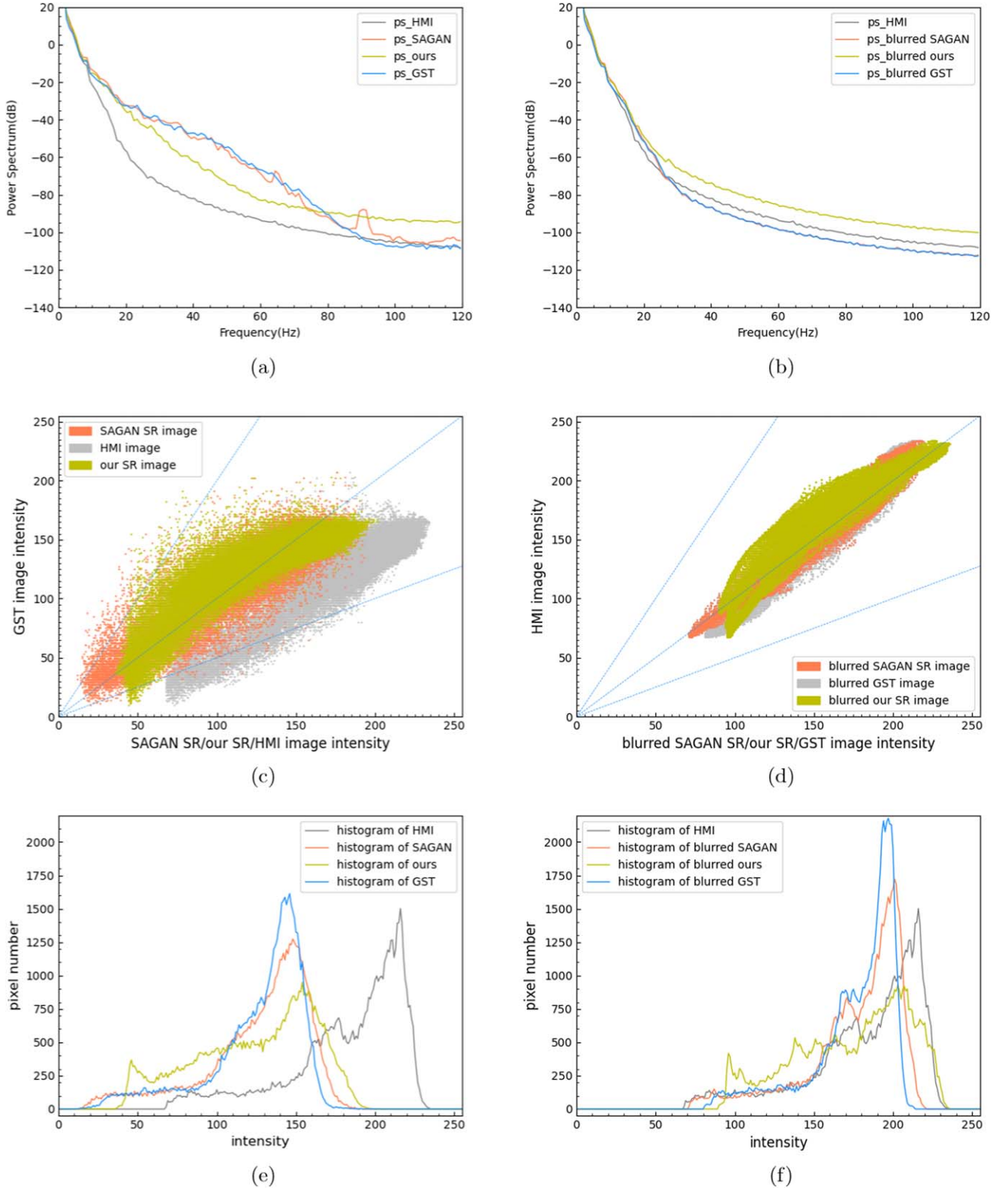


Figure 7. The three rows are the comparison of the power spectrum, scatter plot, and gray histogram of the SR image, GST image, and HMI image, respectively. The dotted lines in the scatter plot represent auxiliary lines with slopes of 0.5, 1, and 2 in turn. The left column compares the HMI image, SR image, and GST image. The right column compares the HMI image, blurred SR image, and blurred GST image.

of ICDDPM preserve and utilize the information of the HMI images as much as possible. In order to compare the consistency of the reconstructed images with the HMI images in terms of structure, contrast, etc., the resolutions of the

reconstructed results from SAGAN and ICDDPM are Gaussian blurred to approximate the resolution of the HMI images. Figure 6(d) shows the similarity of the blurred results of SAGAN and ICDDPM to the HMI image, respectively.

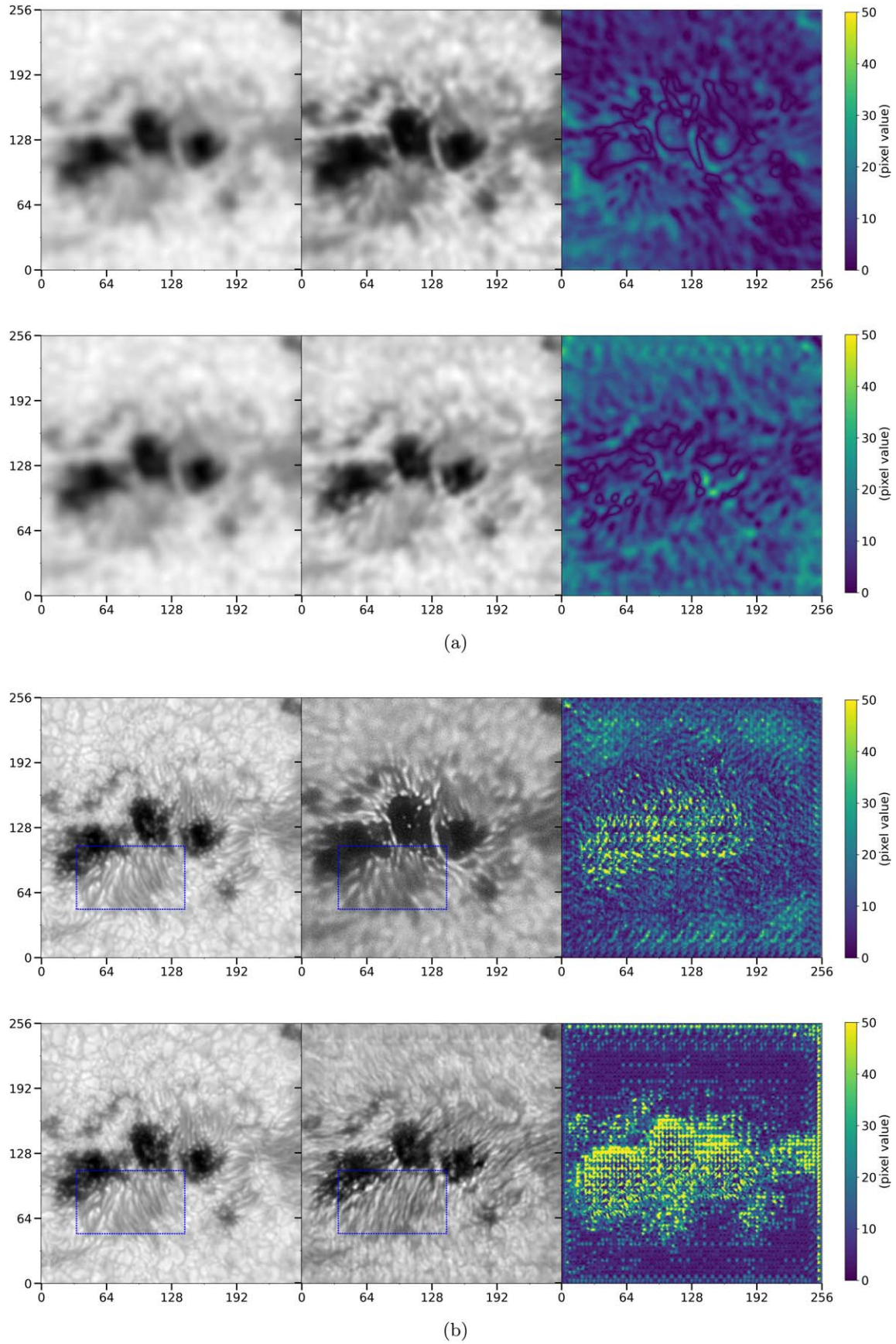


Figure 8. (a) Shows the HMI image, the SR image blurred to the resolution of the HMI image, and the residual image from left to right; (b) shows the GST image, the SR image, and the prediction residual image from left to right. Blue dotted boxes are used to facilitate the comparison of images. The SR image in the first row in (a) and (b) is the result of our model, and the SR image in the second row is the result of SAGAN.

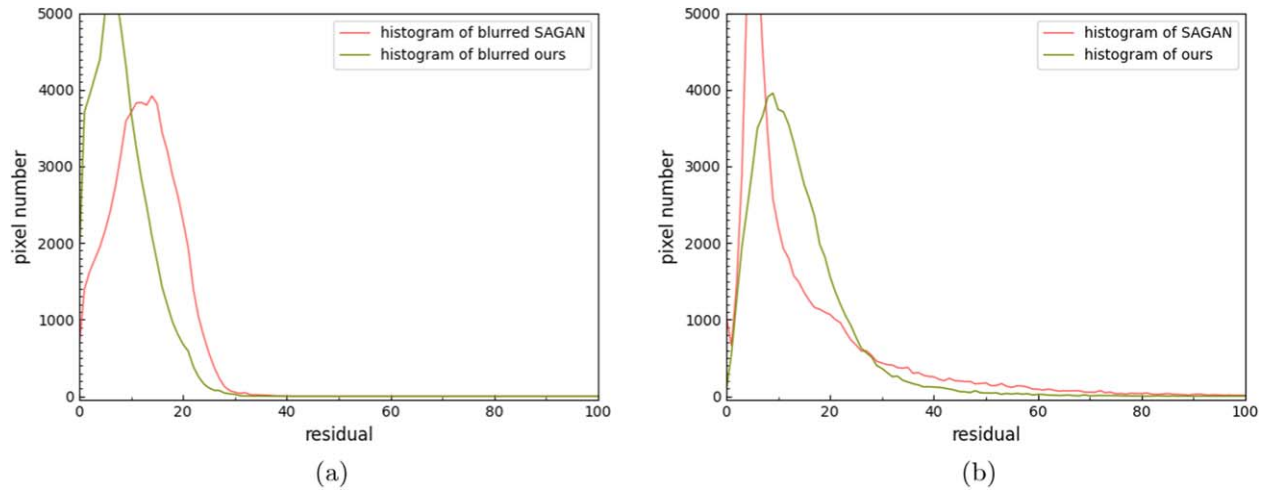


Figure 9. The resolution of the SR image in (a) is blurred to approximate the resolution of the HMI image. (a) The histogram of the residual image of the blurred SR image and HMI image. (b) The histogram of the prediction residual image of the SR image and the GST image. The green and red curves in (a) and (b) represent the residual values obtained from our results and the results of SAGAN, respectively.

Overall, their results are well consistent with the HMI image. The average value of SSIM of the results of ICDDPM is 0.868. This value is 0.007 higher than the results of SAGAN. This may indicate a slight advantage in the consistency of the reconstruction results of ICDDPM and the HMI images.

This paper also plots the azimuthally average power spectrum of the HMI image, GST image, our SR image, and SAGAN SR image. This is used to compare the information content of each frequency in these four images. As shown in Figure 7(a), compared with the HMI image, the mid-to-high frequency of our SR image is improved to a certain extent. This may represent an enhancement of the basic structure of the image. The power spectrum curve of SAGAN fluctuates around 90 Hz, in contrast, the curve of ICDDPM is more stable. This may be because SAGAN introduces some artifacts in the high-frequency information while reconstructing the details, causing the curve to fluctuate. But these artifacts are also blurred when the image is blurred, so there is no fluctuation in the power spectrum curve of the blurred result of SAGAN (Figure 7(b)). In Figures 7(a) and (b), the trends of the curves show that the result of SAGAN is similar to those of the GST image. And the result of ICDDPM is similar to those of the HMI image. This may further illustrate the similarity of the reconstruction results of SAGAN with GST images. And the consistency of reconstruction results of ICDDPM with HMI images. Figure 7(c) is a scatter plot of the SR image and the GST image to compare the pixel values of corresponding points in the two images. The more concentrated the generated points are on a straight line with a slope of 1, the more similar the image is to the corresponding pixels of the GST image. In Figure 7(c), the gray, red, and green dots represent the correlations of the HMI image, the SR image of SAGAN, and the SR image of ICDDPM with the GST image, respectively. Compared to SAGAN, the green dots representing our SR images are more compactly concentrated near the straight line with a slope of 1 in the region with pixel values of $60 \sim 150$. In the regions with pixel values around 50 and 160, the overall slope of the green point is similar to the gray point slope of the HMI image. This may be because our SR image retains the main structural features of the input HMI image while reconstructing certain details. Figure 7(d) represents the similarity between the

blurred image and the HMI image. Our blurred SR image produces green dots with an overall slope close to 1. This represents a strong linear relationship with the HMI image. Figures 7(e) and (f) are the gray histograms of different images, which are used to compare the number of pixels with different grayscale values. It can be seen from Figure 7(e) that the pixel value of the HMI image is mainly concentrated around 200. This reflects that the overall brightness of the HMI image is large, which is not conducive to observation. The results of ICDDPM are more uniform in the intensity distribution of the histogram than those of SAGAN. This means that the overall contrast of the image is enhanced, which is consistent with the information reflected in Figure 6(c). This helps to observe the details of the image more clearly. In Figures 7(e) and (f), fluctuations appear near the left side of the curve of the SR image of ICDDPM. This may be due to the wide borders of the reconstructed granules increasing the number of low-brightness pixels. In terms of the peaks of the curves in Figure 7(f), our result is slightly lower than the HMI image, and the SAGAN result is slightly lower than the GST image. This may be due to the image contrast being altered during the reconstruction of some details. By analyzing Figure 7, the different reconstruction requirements of ICDDPM and SAGAN are further verified. SAGAN focuses on SR reconstruction results with the same rich details as GST images. ICDDPM focuses on reconstructing more details while being faithful to the HMI image.

The task of image SR is an ill-posed problem, and the uncertainty of SR has an impact on the accuracy of the model output (Gitiaux et al. 2019). For scientific applications, in addition to good visual effects, the estimation of errors and uncertainties is also important (Gitiaux et al. 2019). To estimate and compare the reconstruction errors of ICDDPM and SAGAN, the HMI images and the residual images of the SR and GST images are used as input to train additional networks. During testing, the HMI images are used as input, and the prediction residuals between the SR and GST images are output. Taking the images in Figure 3(e) as an example, the residual images obtained from them are shown in Figure 8. Figure 8(a) shows the residuals between the blurred reconstruction results and the HMI image. In the result of ICDDPM,

the differences are concentrated in the lower left corner of the reconstructed image. The contrast of this part of the reconstructed image slightly deviates from the HMI image. It is consistent with the conclusion in Figure 7(f). In the result of SAGAN, the differences are concentrated at the edges of the image. This may be caused by the blurring artifacts. Figure 8(b) shows the residuals between the reconstruction results and the GST image. In the results of ICDDPM, the differences are concentrated at the junction of the umbra and penumbra. This shows that the reconstruction of the junction needs to be strengthened. For SAGAN, the differences are concentrated in the penumbra. This may be because there are more artifacts in this part. For ease of comparison, the histograms of the residual images calculated from the SR reconstruction results in Figure 8 are shown in Figure 9. The range of residual value is [0, 255], and the smaller the value, the smaller the difference. Both curves tend to 0 where the residual value is greater than 100, so the values on the horizontal axis only show [0, 100]. In Figure 9(a), the peaks of the curves of ICDDPM and SAGAN results are approximately at residual values of 5 and 15, respectively. In contrast, the curve of ICDDPM is closer to 0, representing less difference. This may further illustrate the consistency between reconstructed images of ICDDPM and HMI images. In Figure 9(b), the peak value of the curve of the ICDDPM result is around the residual value of 10, which is slightly worse than SAGAN. But in the region of 30 ~ 70, the ICDDPM curve is lower than SAGAN. That is, ICDDPM has fewer pixels with larger residual errors than SAGAN. This is consistent with the information reflected in Figure 8(b). At the same time, this may also mean that ICDDPM has relatively small errors in the reconstruction process.

5. Conclusion

This paper proposes an SR reconstruction model of solar images based on conditional DDPMs. During training, the new auxiliary loss term is added. During reconstruction, the given HMI images are used to construct the initialized images and refine the results of each iterative step. By combining the improvement of the loss function and the reconstruction process, the model finally achieves the goal of reconstructing the HMI images with four times the spatial resolution. We evaluate this model both subjectively and objectively, and compare it with SAGAN. The results show that our model can complete the task of SR reconstruction well while being faithful to the HMI image. This proves the feasibility of our model in the SR reconstruction of solar images.

In the future, our goal is to reconstruct finer textures while maintaining consistency with LR images. Different aspects will be explored, such as in terms of data sets: more advanced image registration algorithms (Liang et al. 2022a) will be used to obtain more accurate image pairs as data sets; in the improvement of the model: more lightweight DDPMs will be studied, trying to reduce the dependence of this model on high-performance hardware, and improve the reconstruction speed while ensuring the quality of reconstruction. Various reconstruction algorithms will be tried to combine effectively, such as DDPMs combined with GAN (Xiao et al. 2021), or combined with solar imaging processes (Jia et al. 2021), to enhance the interpretability of the model while improving the effect of reconstruction. We will also extend this work to SR reconstruction of other solar image data sets to test the universality of the model, such as the H-alpha image observed

by NVST is used as the HR image, the H-alpha full-disk observation data provided by Huairou Solar Observing Station (HSOS) is used as the LR image.

The work is supported by the open project of CAS Key Laboratory of Solar Activity (grant No. KLSA202114), Young Academic Team Leadership Program (grant No. 2022QNYL31), and Ability Enhancement Project for Scientific Research Management (grant No. 2021GLTS15).

ORCID iDs

Wei Song  <https://orcid.org/0000-0002-2324-4302>

References

- Abramenko, V. I., & Yurchyshyn, V. B. 2020, *MNRAS*, **497**, 5405
- Ajabshirizadeh, A., Jouzdani, N. M., & Abbassi, S. 2011, *RAA*, **11**, 491
- Bamba, Y., Kusano, K., Imada, S., & Iida, Y. 2014, *PASJ*, **66**, S16
- Baso, C. D., de la Cruz Rodríguez, J., & Danilovic, S. 2019, *A&A*, **629**, A99
- Baso, C. D., & Ramos, A. A. 2018, *A&A*, **614**, A5
- Baso, C. D., Ramos, A. A., & de la Cruz Rodríguez, J. 2022, *A&A*, **659**, A165
- Chae, J., Park, H.-M., Ahn, K., et al. 2013, *SoPh*, **288**, 1
- Choi, J., Kim, S., Jeong, Y., Gwon, Y., & Yoon, S. 2021, in 2021 IEEE/CVF Int. Conf. on Computer Vision (ICCV) (Piscataway, NJ: IEEE), 14347
- Deng, J., Song, W., Liu, D., et al. 2021, *ApJ*, **923**, 76
- Denker, C., Dineva, E., Balthasar, H., et al. 2018, *SoPh*, **293**, 44
- Dhariwal, P., & Nichol, A. 2021, arXiv:2105.05233
- Engels, S., & Van Geel, B. 2012, *JSWSC*, **2**, A09
- Feulner, G., & Rahmstorf, S. 2010, *GeoRL*, **37**, L05707
- Gitiaux, X., Maloney, S. A., Jungbluth, A., et al. 2019, arXiv:1911.01486
- Goode, P. R., & Cao, W. 2012, *Proc. SPIE*, **8444**, 844403
- He, K., Zhang, X., Ren, S., & Sun, J. 2016, in 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Piscataway, NJ: IEEE), 770
- Ho, J., Jain, A., & Abbeel, P. 2020, *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*, ed. H. Larochelle et al., 6840, <https://proceedings.neurips.cc/paper/2020/hash/4c5bfc8584af0d967f1ab10179ca4b-Abstract.html>
- Jia, P., Huang, Y., Cai, B., & Cai, D. 2019, *ApJL*, **881**, L30
- Liang, B., Chen, X., Li, Z., et al. 2021, *MNRAS*, **505**, 4717
- Jia, P., Wu, X., Yi, H., Cai, B., & Cai, D. 2020, *AJ*, **159**, 183
- Jonas, E., Bobra, M., Shankar, V., Todd Hoeksema, J., & Recht, B. 2018, *SoPh*, **293**, 48
- Kingma, D. P., & Welling, M. 2013, arXiv:1312.6114
- Li, F., Jiang, M., & Yang, L. 2021, in *Pattern Recognition and Computer Vision*, ed. H. Ma et al. (Cham: Springer), 496
- Li, H., Yang, Y., Chang, M., et al. 2022, *Neurocomputing*, **479**, 47
- Liang, B., Chen, X., Yu, L., et al. 2022a, *ApJS*, **261**, 10
- Liang, J., Zeng, H., & Zhang, L. 2022b, in 2022 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR) (Los Alamitos, CA: IEEE Computer Society), 5647
- Liu, H., Jin, Z.-Y., Xiang, Y., & Ji, K. 2022, *RAA*, **22**, 095005
- Löfdahl, M. G. 2002, *Proc. SPIE*, **4792**, 146
- Löfdahl, M. G., Hillberg, T., de la Cruz Rodríguez, J., et al. 2021, *A&A*, **653**, A68
- Lowe, D. G. 2004, *Int. J. Comput. Vis.*, **60**, 91
- Lugmayr, A., Danelljan, M., Romero, A., et al. 2022, arXiv:2201.09865
- Meng, C., Song, Y., Song, J., et al. 2021, arXiv:2108.01073
- Nelson, C., Doyle, J., Erdélyi, R., et al. 2013, *SoPh*, **283**, 307
- Nichol, A. Q., & Dhariwal, P. 2021, *Proc. Machine Learning Research 139*, Proc. of the 38th Int. Conf. on Machine Learning, ed. M. Meila & T. Zhang, (PMLR), 8162, <https://proceedings.mlr.press/v139/nichol21a.html>
- Popowicz, A., Radlak, K., Bernacki, K., & Orlov, V. 2017, *SoPh*, **292**, 187
- Rahman, S., Moon, Y.-J., Park, E., et al. 2020, *ApJL*, **897**, L32
- Ramos, A. A., de la Cruz Rodríguez, J., & Yabar, A. P. 2018, *A&A*, **620**, A73
- Ramos, A. A., & Olsper, N. 2021, *A&A*, **646**, A100
- Rezende, D., & Mohamed, S. 2015, *Proc. Machine Learning Research 37*, Proc. of the 32nd Int. Conf. on Machine Learning, ed. F. Bach & D. Blei, (PMLR), 2256, <https://proceedings.mlr.press/v37/sohl-dickstein15.html>
- Saharia, C., Ho, J., Chan, W., et al. 2022, *ITPAM*, in press
- Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, *SoPh*, **275**, 229
- Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., & Ganguli, S. 2015, *Proc. Machine Learning Research 37*, Proc. of the 32nd Int. Conf. on Machine

- Learning, ed. F. Bach & D. Blei, (PMLR), 1530, <http://proceedings.mlr.press/v37/rezende15.html>
- Song, J., Meng, C., & Ermon, S. 2020, arXiv:2010.02502
- Song, Y., & Ermon, S. 2019, Advances in Neural Information Processing Systems 32 (NeurIPS 2019), ed. H. Wallach et al., <https://papers.nips.cc/paper/2019/hash/3001ef257407d5a371a96dcd947c7d93-Abstract.html>
- Stengel, K., Glaws, A., Hettinger, D., & King, R. N. 2020, *PNAS*, **117**, 16805
- Svalgaard, L. 2013, *JSWSC*, **3**, A24
- Tian, F., Budavári, T., Basu, A., & Lubow, S. H. L. W. R. 2019, *AJ*, **158**, 191
- Usoskin, I. G. 2017, *LRSP*, **14**, 1
- Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., et al. 2014, *PeerJ*, **2**, e453
- Vaswani, A., Shazeer, N., Parmar, N., et al. 2017, Advances in Neural Information Processing Systems 30 (NIPS 2017), ed. I. Guyon, <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>
- Wang, H., & Liu, C. 2019, *FrASS*, **6**, 18
- Wang, Q., Zhou, H., Li, G., & Guo, J. 2022, *Appl. Sci.*, **12**, 6067
- Wang, S., Chen, Q., He, C., et al. 2021, *A&A*, **652**, A50
- Whang, J., Delbracio, M., Talebi, H., et al. 2021, arXiv:2112.02475
- Wu, H., Qu, Y., Lin, S., et al. 2021, in 2021 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR) (Piscataway, NJ: IEEE), 10546
- Wu, Y., & He, K. 2018, *Int. J. Comput. Vis.*, **128**, 742
- Xiao, Z., Kreis, K., & Vahdat, A. 2021, arXiv:2112.07804
- Xie, C., Zhu, H., & Fei, Y. 2022, *IET Image Process.*, **16**, 273