

Advances in Astronomy

# Big Data Processing and Modeling in Solar Physics

Lead Guest Editor: Xin Huang

Guest Editors: Ilya Usoskin, Liyun Zhang, and Huaning Wang





---

# **Big Data Processing and Modeling in Solar Physics**

Advances in Astronomy

---

## **Big Data Processing and Modeling in Solar Physics**

Lead Guest Editor: Xin Huang

Guest Editors: Ilya Usoskin, Liyun Zhang, and Huaning Wang



---

Copyright © 2020 Hindawi Limited. All rights reserved.

This is a special issue published in "Advances in Astronomy." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## **Chief Editor**

Josep M. Trigo-Rodríguez, Spain

---

## **Editorial Board**

Lorenzo Amati, Italy  
Miguel De Avillez, Portugal  
Damien A. Easson, USA  
Dean Hines, USA  
John Hughes, USA  
Wing-Huen Ip, Taiwan  
Geza Kovacs, Hungary  
Michael Kueppers, Spain  
Charalampos C. Moustakidis, Greece  
Zdzislaw E. Musielak, USA  
Valery Nakariakov, United Kingdom  
George Pavlov, USA  
Alexei S. Pozanenko, Russia  
Somak Raychaudhury, India  
Wenwu Tian, China  
Roberto Turolla, Italy  
Yue Wang, China

# Contents

---

## **Big Data Processing and Modeling in Solar Physics**

X. Huang , I. Usoskin, L. Y. Zhang, and H. N. Wang

Editorial (1 page), Article ID 6967925, Volume 2020 (2020)

## **Mask-Pix2Pix Network for Overexposure Region Recovery of Solar Image**

Dong Zhao , Long Xu , Linjie Chen, Yihua Yan, and Ling-Yu Duan

Research Article (10 pages), Article ID 5343254, Volume 2019 (2019)

## **Intelligent Recognition of Time Stamp Characters in Solar Scanned Images from Film**

Jiafeng Zhang , Guangzhong Lin, Shuguang Zeng , Sheng Zheng, Xiao Yang, Ganghua Lin, Xiangyun Zeng, and Haimin Wang 

Research Article (9 pages), Article ID 6565379, Volume 2019 (2019)

## **Full-Disk Solar Flare Forecasting Model Based on Data Mining Method**

Rong Li  and Yong Du

Research Article (6 pages), Article ID 5190353, Volume 2019 (2019)

## **Deep Learning for Automatic Recognition of Magnetic Type in Sunspot Groups**

Yuanhui Fang , Yanmei Cui , and Xianzhi Ao

Research Article (10 pages), Article ID 9196234, Volume 2019 (2019)

## **Effect of Ap-Index of Geomagnetic Activity on S&P 500 Stock Market Return**

Lifang Peng, Ning Li , and Jingwen Pan

Research Article (8 pages), Article ID 2748062, Volume 2019 (2019)

## **Single and Multiwavelength Detection of Coronal Dimming and Coronal Wave Using Faster R-CNN**

Zongxia Xie and Chunyang Ji 

Research Article (9 pages), Article ID 7821025, Volume 2019 (2019)

## **The Mid-Term Forecast Method of F10.7 Based on Extreme Ultraviolet Images**

L. Lei , Q. Zhong , J. Wang, L. Shi, and S. Liu

Research Article (14 pages), Article ID 5604092, Volume 2019 (2019)

## Editorial

# Big Data Processing and Modeling in Solar Physics

**X. Huang** <sup>1</sup>, **I. Usoskin**,<sup>2</sup> **L. Y. Zhang**,<sup>2</sup> and **H. N. Wang**<sup>1</sup>

<sup>1</sup>Key Laboratory of Solar Activity, National Astronomical Observatories of Chinese Academy of Sciences, Beijing, China

<sup>2</sup>Space Physics Unit and Sodankylä Geophysical Observatory, University of Oulu, Oulu, Finland

Correspondence should be addressed to X. Huang; xhuang@bao.ac.cn

Received 25 February 2020; Accepted 25 February 2020; Published 17 March 2020

Copyright © 2020 X. Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Sun is the energy source of the Earth. The electromagnetic environment of the Earth is affected by solar activity, and the impact of violent activity bursts can reach the Earth within eight minutes. Hence the detection, recognition, and prediction of solar activity are essential.

The physical mechanisms of solar activity bursts are not yet completely clear. However, a large number of data have been accumulated and solar observation instruments can record the multiwavelength imaging data every day with high cadence. In order to cope with the rapidly growing amount of solar data, there is an increasing need for automatic detection and prediction technologies.

This special issue is focused on solar data mining technology. We invited authors to contribute with original research articles in this special issue. Eleven original research manuscripts have been received. After the peer-reviewed process, seven of them were accepted for publications. Therein, three papers focused on the detection and recognition of regions of interest in the solar images, two papers presented research on the short-term and midterm solar activity prediction, respectively, and one paper discussed the influence of solar activity on economic activities. From these articles, we can find that the machine learning methods, especially the deep learning methods, play an important role in solar activity monitoring and prediction. Finally, we hope that researchers will find this special issue useful.

## Conflicts of Interest

The editors declare that they have no conflicts of interest regarding the publication of this special issue.

## Acknowledgments

We sincerely thank the authors contributing to the special issue and referees thoroughly reviewing the papers. The special issue also benefits from the members of the editorial board of Advances in Astronomy.

X. Huang  
I. Usoskin  
L. Y. Zhang  
H. N. Wang

## Research Article

# Mask-Pix2Pix Network for Overexposure Region Recovery of Solar Image

Dong Zhao <sup>1,2</sup>, Long Xu <sup>1</sup>, Linjie Chen,<sup>1</sup> Yihua Yan,<sup>1</sup> and Ling-Yu Duan<sup>3</sup>

<sup>1</sup>Key Laboratory of Solar Activity, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China

<sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup>National Engineering Lab for Video Technology, Peking University, Beijing 100871, China

Correspondence should be addressed to Long Xu; [lxu@nao.cas.cn](mailto:lxu@nao.cas.cn)

Received 1 May 2019; Accepted 10 August 2019; Published 5 September 2019

Academic Editor: Zdzislaw E. Musielak

Copyright © 2019 Dong Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Overexposure may happen for imaging of solar observation as extremely violet solar bursts occur, which means that signal intensity goes beyond the dynamic range of imaging system of a telescope, resulting in loss of signal. For example, during solar flare, Atmospheric Imaging Assembly (AIA) of Solar Dynamics Observatory (SDO) often records overexposed images/videos, resulting loss of fine structures of solar flare. This paper makes effort to retrieve/recover missing information of overexposure by exploiting deep learning for its powerful nonlinear representation which makes it widely used in image reconstruction/restoration. First, a new model, namely, mask-Pix2Pix network, is proposed for overexposure recovery. It is built on a well-known Pix2Pix network of conditional generative adversarial network (cGAN). In addition, a hybrid loss function, including an adversarial loss, a masked L1 loss and a edge mass loss/smoothness, are integrated together for addressing challenges of overexposure relative to conventional image restoration. Moreover, a new database of overexposure is established for training the proposed model. Extensive experimental results demonstrate that the proposed mask-Pix2Pix network can well recover missing information of overexposure and outperforms the state of the arts originally designed for image reconstruction tasks.

## 1. Introduction

Solar imaging can provide us more information than one-dimensional solar flux about solar activities, promoting our abilities of probing the mystery of the Sun. Especially, the instruments onboard satellites have more advantages than ground-based instruments, due to good seeing and non-occlusion. During the past few decades, number of satellites carrying solar probing instruments have been launched, including Yohkoh [1], SOHO [2–4], TRACE [5], STEREO [6, 7], Hinode [8], and SDO [9–12]. These space instruments can provide us solar observation with unprecedented time cadence, high dynamic range, and high spatial/frequency resolution. For example, Atmospheric Imaging Assembly (AIA) instrument onboard SDO can image solar atmosphere in 12 wavelengths every 12 seconds, with spatial resolution of  $4096 \times 4096$ .

Nowadays, satellite carrying solar instruments can image the complete process of a solar activity with unprecedented high time cadence and high spatial resolution. These

recorded images and videos provide the scientists the opportunity for uncovering the nature of solar activities, such as sunspot, filament, coronal loop, flare, or coronal mass emission (CME). However, for some extremely violet solar bursts, the intensity of signal may go beyond of the threshold of a telescope, leading to overexposure; therefore, information is missed in the captured images/videos. Although there were automatic exposure-control algorithms which were implemented in case of flares, controlling of exposure time to reduce overexposure, for example, AIA avoids overexposure by alternating between low and long exposures, they cannot guarantee complete avoidance of overexposure. In addition, the time cadence would increase, so temporal resolution is compromised. For addressing overexposure, an iterative algorithm by employing the PRiL approximation and EM algorithm is proposed in [13]. Different from [13], we solve this issue via deep learning.

Recently, taking advantages of the powerful capability of deep learning in its nonlinear representation, lots of

traditional image processing/reconstruction problems achieved new breakthroughs, such as image noising, superresolution, image inpainting. Therefore, in this work, we decide to restore overexposure region (OER) by taking advantage of deep learning. Inspired by image inpainting, our task can be recast into a similar optimization problem of image inpainting. Both of them are proposed to recover missing region in an image using surrounding pixels according to the continuity of signal. However, for our task, there are at least three challenges relative to traditional image inpainting. Firstly, the fidelity of signal is required, not just look-alikes, since the data concerned in our task are used for scientific purpose, for example, computing physical parameters. Secondly, larger OER is concerned in our task comparing to image painting. The last, more irregular shapes of OER result in complicated boundaries connecting OERs and nonOERs/normal regions. These boundaries should be suppressed to prevent unnatural change of pixel value.

The existing image inpainting approaches can be divided into three classes: diffusion-based, patch-based, and learning based. From our experience, neither of diffusion-based and patch-based methods can contribute effective results for OER recovery. The diffusion-based methods [14,15] are highly restricted to locally available information, so they fail to recover meaningful structures of large missing regions, like OERs in our tasks. Patch-based methods [16–19] assume that missing region (OER) can find its most similar/relevant patches in the given image. This assumption, however, does not hold for OERs. The existing learning-based inpainting methods are not competent in our task for two reasons. First, some of the learning-based methods, such as [20–23] and [24], trained their networks with fixed shapes and locations of missing regions, while the OERs commonly exhibit irregular shape and random location. Second, despite several studies for irregular missing region in image inpainting, such as [25–28], they focus on generating visually coherent completion or producing semantically plausible results. However, our task aims to restore missing OER with high fidelity, not just a visually plausible result.

To address the aforementioned challenges, we propose a learning-based model, namely, mask-Pix2Pix network, to estimate the OER of solar image. Our network is built on the Pix2Pix [29], which is a general network for image-to-image translation using a conditional generative adversarial network (cGAN) [30]. We adopt the similar architectures with the Pix2Pix: U-net [31] in generator and PatchGAN [30] in discriminator. The main improvements and contributions of this paper are as follows: (1) Unlike the conventional Pix2Pix, our network utilizes the Convolution-SwitchNorm-LReLU/ReLU [32] modules (LReLU for encoder and ReLU for decoder) rather than the Convolution-BatchNorm-ReLU [33] ones. The former (i.e., switchable normalization) can switch between BatchNorm [33], LayerNorm [34], and InstanceNorm [35] by learning their importance weights in an end-to-end manner. The improved architecture of our model boosts the robustness of the network. (2) Our objective function contains an adversarial cGAN loss, a masked L1 loss, and an edge mask loss/smoothness. The adversarial cGAN loss can capture the full entropy of the conditional

distributions they model and thereby produce highly realistic textures. The masked L1 loss calculates the L1 loss only in masked regions (OERs), enforcing correctness at low frequencies which guarantees restoration of high fidelity for OERs. The edge mask loss is used for smoothing edges of OERs and suppressing edge artifacts in final restored image. (3) Additionally, a new database of overexposure is established for training and testing the proposed mask-Pix2Pix network, where 13700 images are collected from the Large-scale Solar Dynamics Observatory image database (LSDO) [36].

## 2. Background

*2.1. Image Inpainting.* Image inpainting is the process of reconstructing lost or deteriorated parts of an image. Nowadays, it still remains theoretically and computationally challenging. It is highly ill-posed. Existing image inpainting algorithms can be categorized into three groups, diffusion-based [14, 15], patch-based [16–19], and learning-based. The former two typically use differential operators or patch similarity to propagate information from normal regions to missing regions. They work well for stationary textures. However, diffusion-based ones are highly restricted to locally available information, so they usually fail to recover meaningful structures of large missing regions. Patch-based ones, such as Planar Structure Guidance (PSG) [18] and Statistics of Similar Patches (SSPs) [19], assume that missing regions can find their most similar/relevant patches from normal regions of the given input image. This assumption does not hold for our task where OER probably does not have similar patches.

Recently, deep learning-based approaches have emerged as a promising paradigm for image inpainting. A significant advantage of them over conventional ones lies in their abilities to learn and understand semantics of images with complicated scenes. Context-Encoder (CE) [21] introduces a deep adversarial model to predict plausible structures with combined 2 pixel-wise reconstruction loss and adversarial loss. Although the adversarial loss improves the inpainting quality, the results exhibit blurriness and contain notable artifacts. To overcome the limitations of the CE [21] method, Iizuka et al. [23] propose a novel architecture with three networks: the completion network (using dilated convolutions [20]), the global, and local context discriminators. However, the training model with two discriminators is time-consuming, and the output heavily relies on Poisson image blending [37] as a postprocessing. Multiscale neural patch synthesis (MNPS) [22] approach contains a content network to learn the semantics and global structure of the image and a texture network to generate fine-detailed result by employing a pretrained VGG-19 [38] network. Liu et al. [25] introduced a partial convolution (PC) model for irregularly shaped hole inpainting with an automatic mask updating step to reduce artifacts. In the PC [25] method, the convolution is masked and renormalized to be conditioned on only valid pixels. Both of the methods of MNPS [22] and PC [25] add the TV loss term to encourage smoothness on the 1-pixel dilation of the hole boundaries. Therefore, they

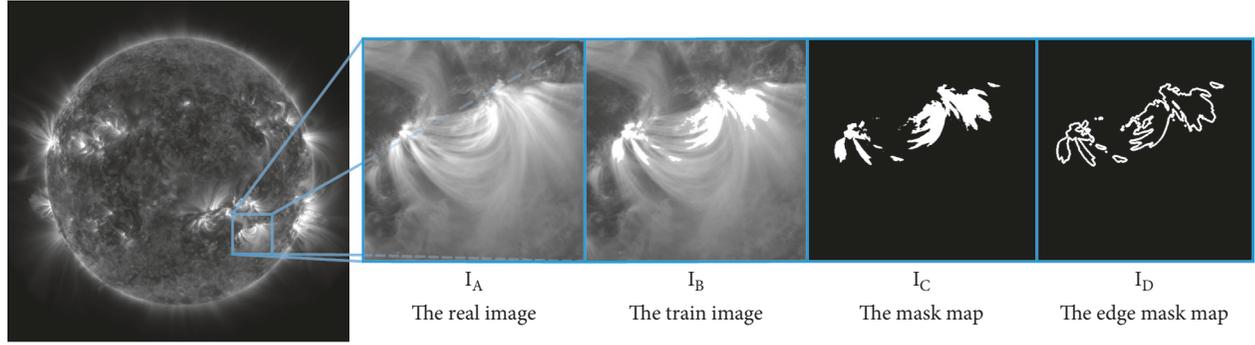


FIGURE 1: Overexposure Region (OER) database built on the LSDO.

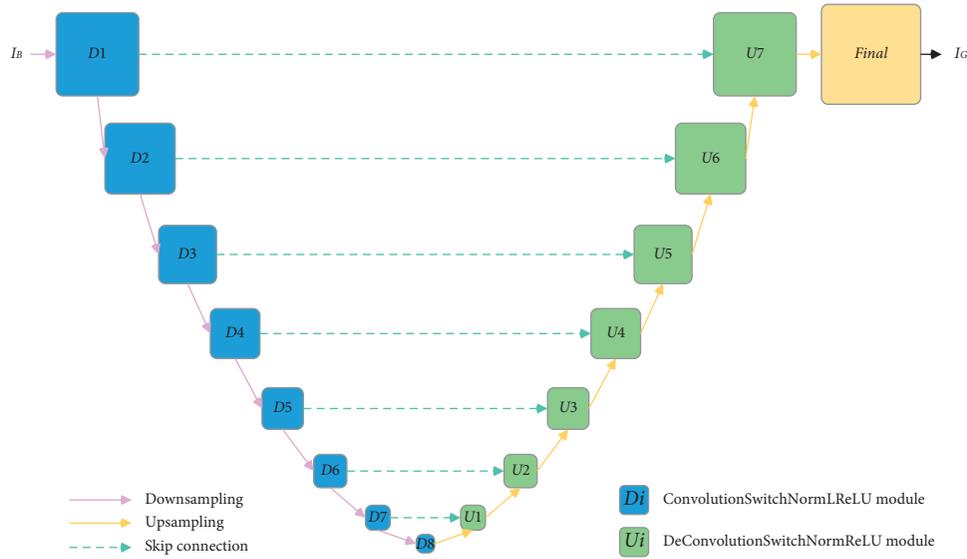


FIGURE 2: Architecture of the generator.

avoid requiring postprocessing steps (such as Poisson image blending operation) to enforce texture coherency near the hole boundaries.

**2.2. Image-to-Image Translation with Pix2Pix.** Pix2Pix [29] utilizes a conditional generative adversarial network (cGAN) [30] to achieve the target of image-to-image translation. Instead of using a conventional encoder-decoder, the generator in Pix2Pix employs an U-Net [31] architecture, in which the encoder layers and decoder layers are directly connected by “skip connection.” Since the skip connection can shuttle the low-level information (which are commonly shared between the input and output images) across the bottleneck of the encoder-decoder net, it effectively improves the performance of the image-to-image translation. The synthesized images of the generator should not be distinguished from the real one by an adversarially trained discriminator. In order to model high frequencies, the discriminator employs a convolutional PatchGAN classifier, i.e., only the structure at certain scale of patches is penalized. By this method, the PatchGAN is taken as a form of texture loss. To summarise, there are at least three factors that enable the Pix2Pix network outperform past works: the U-Net

architecture for the generator, the convolutional PatchGAN classifier for discriminator, and the cGANs for network training.

### 3. Proposed Method

**3.1. Problem Description.** The overexposure problem is simply explained in Figure 1, where ground-truth image  $I_A$ , overexposure image  $I_B$ , binary mask map  $I_C$ , and edge mask map  $I_D$  are concerned. The mission of OER recovery is to recover the region labeled by  $I_C$  in  $I_B$ , and outside region of  $I_C$  keeps unchanged. Given OER  $\Omega_M$  and nonOER  $\Omega_{\bar{M}}$ , we can get  $\Omega_M = I_B \odot I_C$  and  $\Omega_{\bar{M}} = I_B \odot (1 - I_C)$ , where  $\odot$  is the element-wise product operator. Inspired by image inpainting, the neural network, GAN, can be employed to retrieve the missing region of  $I_B$  (i.e.,  $\Omega_M$ ). In a GAN, a generator  $G$  is trained on the pairs of ground truth and degraded ones. Then, the generator is applied to a degraded image to output a repaired one, i.e.,  $I_G = G(I_B)$ .

For our task,  $I_G$  should meet the following criteria:

- The restored image should have realistic textures as far as possible, i.e., visually coherent and semantically plausible relative to  $I_A$

TABLE 1: Architecture of the generator.

Layers	Architecture of generator	Output size
Input	$I_A, I_B, I_C, I_D$	$(256 \times 256 \times 1)$
D1	Conv. $(4 \times 4 \times 64)$ , LReLU	$(128 \times 128 \times 64)$
D2	Conv. $(4 \times 4 \times 128)$ , SwitchNorm, LReLU	$(64 \times 64 \times 128)$
D3	Conv. $(4 \times 4 \times 256)$ , SwitchNorm, LReLU	$(32 \times 32 \times 256)$
D4	Conv. $(4 \times 4 \times 512)$ , SwitchNorm, LReLU, dropout	$(16 \times 16 \times 512)$
D5	Conv. $(4 \times 4 \times 512)$ , SwitchNorm, LReLU, dropout	$(8 \times 8 \times 512)$
D6	Conv. $(4 \times 4 \times 512)$ , SwitchNorm, LReLU, dropout	$(4 \times 4 \times 512)$
D7	Conv. $(4 \times 4 \times 512)$ , SwitchNorm, LReLU, dropout	$(2 \times 2 \times 512)$
D8	Conv. $(4 \times 4 \times 512)$ , LReLU, dropout	$(1 \times 1 \times 512)$
U1	Concatenate (D8, D7), DeConv. $(4 \times 4 \times 512)$ , SwitchNorm, ReLU, dropout	$(2 \times 2 \times 512)$
U2	Concatenate (U1, D6), DeConv. $(4 \times 4 \times 512)$ , SwitchNorm, ReLU, dropout	$(4 \times 4 \times 512)$
U3	Concatenate (U2, D5), DeConv. $(4 \times 4 \times 512)$ , SwitchNorm, ReLU, dropout	$(8 \times 8 \times 512)$
U4	Concatenate (U3, D4), DeConv. $(4 \times 4 \times 512)$ , SwitchNorm, ReLU, dropout	$(16 \times 16 \times 512)$
U5	Concatenate (U4, D3), DeConv. $(4 \times 4 \times 256)$ , SwitchNorm, ReLU	$(32 \times 32 \times 256)$
U6	Concatenate (U5, D2), DeConv. $(4 \times 4 \times 128)$ , SwitchNorm, ReLU	$(64 \times 64 \times 128)$
U7	Concatenate (U6, D1), DeConv. $(4 \times 4 \times 64)$ , SwitchNorm, ReLU	$(128 \times 128 \times 64)$
Final	Upsample $(4 \times 4 \times 1)$ , ZeroPad, Conv. $(4 \times 4 \times 1)$ , tanh	$(256 \times 256 \times 1)$
Output	$I_G$	$(256 \times 256 \times 1)$

TABLE 2: Architecture of the discriminator.

Layers	Architecture of discriminator	Output size
Input	$[I_A I_B]$ or $[I_G^* I_B]$	$(256 \times 256 \times 2)$
C1	Conv. $(4 \times 4 \times 64)$ , LReLU	$(128 \times 128 \times 64)$
C2	Conv. $(4 \times 4 \times 128)$ , SwitchNorm, LReLU	$(64 \times 64 \times 128)$
C3	Conv. $(4 \times 4 \times 256)$ , SwitchNorm, LReLU	$(32 \times 32 \times 256)$
C4	Conv. $(4 \times 4 \times 512)$ , SwitchNorm, LReLU	$(16 \times 16 \times 512)$
C5	ZeroPad2d, Conv. $(4 \times 4 \times 1)$	$(16 \times 16 \times 1)$
Output	Real or fake matrix $16 \times 16$	

- (b) The restored region  $I_G \odot I_C$  should have high fidelity relative to corresponding regions in ground-truth  $I_A$
- (c) It should be avoided artificial edges ( $I_C$ ) surrounding OERs, i.e., the boundary between OERs and normal regions should transit smoothly.

To achieve (3.1) and (3.1), conventional Pix2Pix network is employed and modified by adding a mask L1 loss term. To address (3.1), an edge mask loss/smoothness term is introduced, where ground-truth edge mask  $I_D$  is given by the edge of mask map  $I_C$  as shown in Figure 1.

**3.2. Solar OER Recovery Database.** For training our model, a database for overexposure is established. The raw data is cropped from LSDO [36]. The LSDO database records consist of three different parts: event records, corresponding

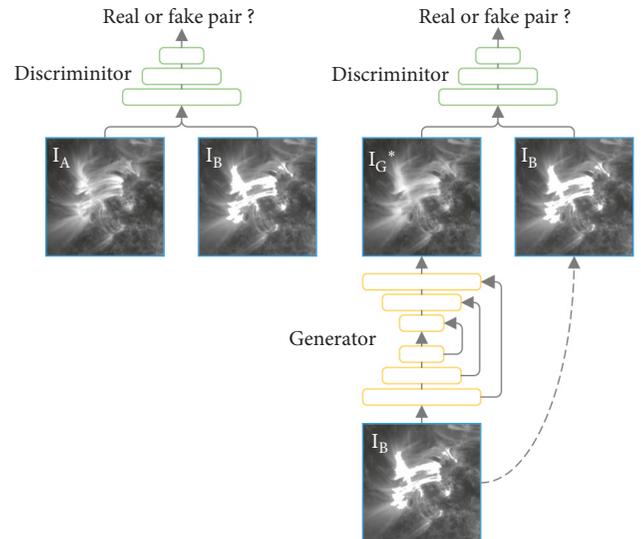


FIGURE 3: Training on adversarial cGAN loss.

images, and extracted image parameters. Event records encompass the list of solar events with generated and extracted attributes. Among these attributes, the bounding box attribute consists of four coordinate pixel values. LSDO has preprocessed all original polygon values of event records from their Helioprojective Coordinates (HPC) solar coordinate system into pixels. This preprocessing step is vital to make polygons compatible with the high-resolution images in the database. More information in detail of the LSDO can be referenced from the literature [36].

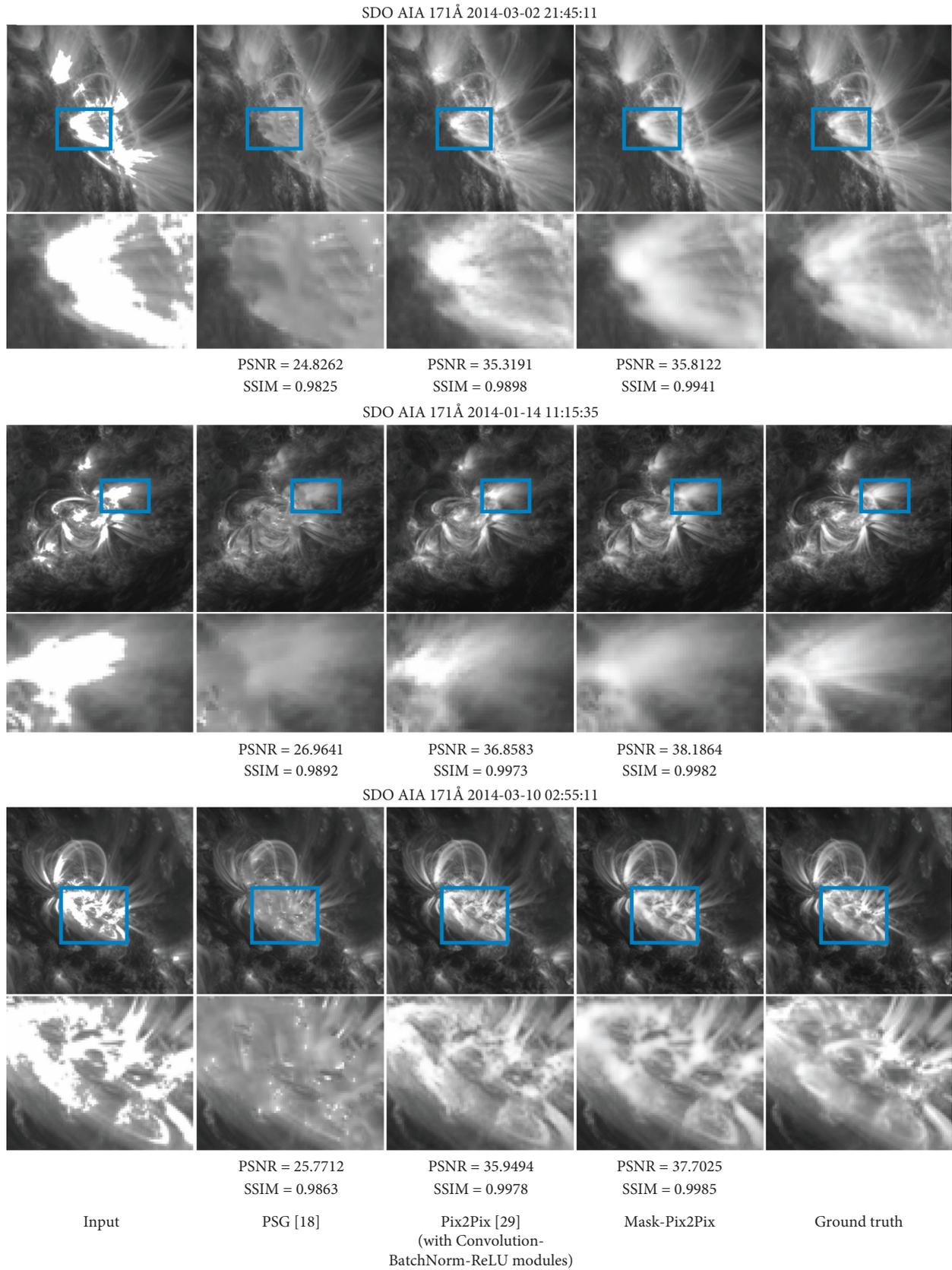


FIGURE 4: Comparisons between the proposed one and two state-of-the-art methods.

TABLE 3: Average PSNR and SSIM comparisons between our model and other two benchmarks.

Methods	Loss functions	Modules	Average PSNR	Average SSIM
PSG [18]			24.6219	0.9810
Pix2Pix [29]	$\mathcal{L}_{\text{cGAN}} + \lambda_1 \mathcal{L}_1^{\text{[a]}}$	Convolution-BatchNorm-ReLU	34.7763	0.9891
Pix2Pix	$\mathcal{L}_{\text{cGAN}} + \lambda_1 \mathcal{L}_1^{\text{[a]}}$	Convolution-SwitchNorm-LReLU/ReLU	35.9939	0.9918
mask-Pix2Pix	$\mathcal{L}_{\text{cGAN}} + \lambda_1 \mathcal{L}_1^{\text{m[a]}}$	Convolution-SwitchNorm-LReLU/ReLU	37.4334	0.9937
mask-Pix2Pix	$\mathcal{L}_{\text{cGAN}} + \lambda_1 \mathcal{L}_1^{\text{m}} + \lambda_2 \mathcal{L}_1^{\text{[b]}}$	Convolution-SwitchNorm-LReLU/ReLU	<b>39.6931</b>	<b>0.9985</b>

<sup>(a)</sup> $\lambda_1 = 0.1$ , <sup>(b)</sup> $\lambda_1 = \lambda_2 = 0.1$ .

In this work, ground-truth  $I_A$  is cropped from Active Region (AR) of images in LSDO. The database is established as following:

- (1) According to polygon values of AR in LSDO, corresponding AR square regions are cropped from nonover-exposed AIA/SDO images, we get  $I_A$ .
- (2) Scale  $I_A$  into the same size of  $(256 \times 256)$ , and label OERs in  $I_A$  by a threshold segmentation: pixel values (from 0 to 1) over a threshold (e.g. 0.8) are set to be 1. Thereby, we get  $I_B$ .
- (3) Label OERs ( $\Omega_M$ ) in  $I_B$  as mask map  $I_C$ , and dilate  $I_C$  with circular kernel (the radius is 3). The dilated mask map  $I_C$  is used to calculate the masked L1 loss in subsection 3.3.
- (4) Extract the edge of  $I_C$ , with edge width of 5 pixel. This edge mask map  $I_D$  accounts for edge mask loss/smoothness in subsection 3.3.

Thereby, each sample in database contains four parts: real image  $I_A$ , overexposed (fake) image  $I_B$ , mask map  $I_C$ , and edge mask map  $I_D$ .

**3.3. Network Architecture.** The proposed mask-Pix2Pix network is composed of a discriminator and a generator. The generator employs a U-Net architecture as demonstrated in Figure 2. The U-Net is named after its shape which looks like a ‘‘U.’’ It consists of an encoder of 8 layers, and a decoder of 8 layers. The parameters of each layer are explained in Table 1 in detail. In addition, skip connections are added between encoder and decoder at the same layer as shown in Figure 2 with dotted lines. Each skip connection simply concatenates feature map of encoder with that of decoder at the same layer (e.g.,  $D_l$  and  $U_{n-l}$  at the  $l$ -th layer in Figure 2). This cross-layer connection can reduce semantic gap between corresponding layers of encoder and decoder since they are too far away in U-Net structure. The discriminator is a PatchGAN network, the structure of which is explained in Table 2. In a GAN framework, the discriminator is to judge ‘‘fake’’ instances from ‘‘real’’ ones. In our work, the output of the discriminator is a  $16 \times 16$  image, each pixel value of which ranges from 0 to 1, for measuring how real is the output. In addition, we adopt Convolution-SwitchNorm-LReLU/ReLU [32] instead of Convolution-BatchNorm-ReLU [33] of conventional Pix2Pix [29] in the proposed model. The former was proved to be more robust.

**3.4. Loss Functions.** It is extremely vital to design a loss function for modelings by using machine learning. Our task is to recover missing signal of OERs in an image, which mostly concerns fidelity of reconstructed signal, natural transition of pixel value. It has something in common with image inpainting, but not the same. Therefore, a new hybrid loss function is designed for OER recovery task as follows.

Given ground-truth image  $I_A$ , degraded image  $I_B$ , initial binary mask  $I_C$  (1 for OERs), and  $I_C$ 's edge map  $I_D$ , as demonstrated in Figure 1, the generator outputs  $I_G = G(I_B)$ . For training model, a hybrid loss function is defined, consisting of three components: (1) adversarial cGAN loss for high fidelity of reconstructed image  $I_G$  relative to ground-truth  $I_A$ , which is defined as

$$\mathcal{L}_{\text{cGAN}}(G, D) = \mathbb{E}[\log D(I_B, I_A)] + \mathbb{E}[\log(1 - D(I_B, G(I_B)))], \quad (1)$$

where  $G$  and  $D$  represents the generator and discriminator, respectively, the objective of  $G$  is to minimize  $D(I_B, G(I_B))$  so that  $G(I_B)$  is more like a real instead of fake image, while  $D$  pursues the maximum of  $D(I_B, I_A)$  to distinguish ‘‘fake’’ and ‘‘real’’ as far as possible. Figure 3 illustrates the process of the adversarial cGAN training, where the discriminator excuses two tasks. The left one is to discriminate if  $I_A$  and  $I_B$  are a pair of ‘‘real’’ and ‘‘fake’’. The right other is to discriminate  $G(I_B)$  is a ‘‘real’’ or ‘‘fake,’’ where the generator  $G$  is trained to produce fakes for deceiving the discriminator  $D$ , while  $D$  is trained to identify ‘‘fake’’ as far as possible.

- (2) L1 loss of masked region  $I_B \odot I_C$  relative to  $I_A \odot I_C$ , targeting for accurate reconstruction of OER, which is defined as:

$$\mathcal{L}_1^{\text{m}}(G) = \mathbb{E}[\| (I_A - G(I_B)) \odot I_C \|_1]. \quad (2)$$

- (3) L1 loss of only edge mask, which could make the edges of OERs smooth to prevent artificial edges connecting OERs and nonOERs. It is defined as

$$\mathcal{L}_1^{\text{e}}(G) = \mathbb{E}[\| (I_A - G(I_B)) \odot I_D \|_1]. \quad (3)$$

In conclusion, the final optimization objective is given by  $G^* = \arg \min_G \max_D \mathcal{L}_{\text{cGAN}}(G, D) + \lambda_1 \mathcal{L}_1^{\text{m}}(G) + \lambda_2 \mathcal{L}_1^{\text{e}}(G)$ , (4)

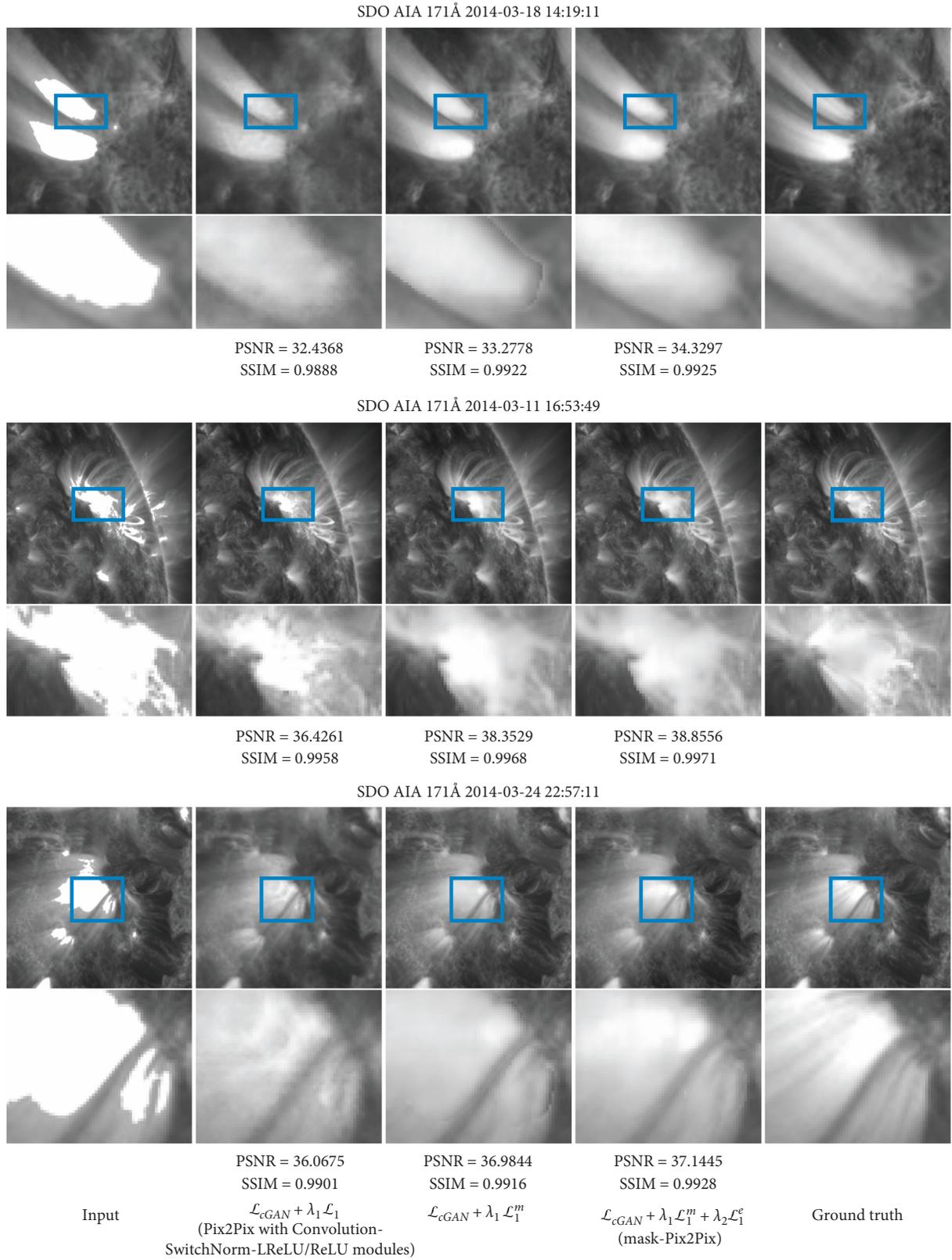


FIGURE 5: Evaluation on the hybrid loss function (contribution of each component is evaluated).

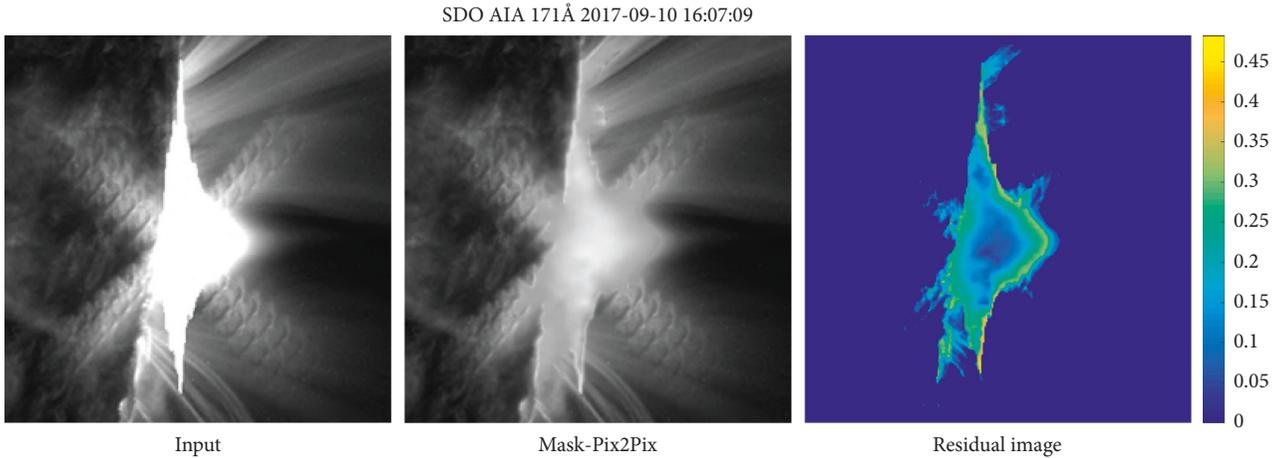


FIGURE 6: Testing on the real image.

where  $\lambda_1$  and  $\lambda_2$  are the weights for combining the above three loss components. They are set to 0.1 in our experience.

#### 4. Experimental Results

To evaluate the proposed mask-Pix2Pix model, it is compared with other state-of-the-art image inpainting algorithms. In addition, the contribution of each factor of the loss function to the overall performance is evaluated. These two experiments are carried out on our synthetic over-exposure images. Then, the pretrained mask-Pix2Pix model is evaluated on real solar images.

**4.1. Implementation Details.** We apply our model to over-exposure training database which are artificially built from LSDO as described in subsection 3.2. For training, the input image size is  $256 \times 256$  and the mini-batch size is set to 16. We train our model with ADAM optimizer [39] by setting  $\beta_1 = 0.5$ ,  $\beta_2 = 0.999$ , and  $\epsilon = 10^{-8}$ . The initial learning rate is initialized to 0.0002 and then reduced to half every 100 epochs. We apply PyTorch on an NVIDIA Tesla P100 for model training.

**4.2. Comparisons with State-of-the-Art Approaches.** We compare the proposed Mask-Pix2Pix with a patch-based Planar Structure Guidance (PSG) [18] and a learning-based Pix2Pix network [29]. Figure 4 shows the comparisons on 3 samples, where it can be found that the proposed Mask-Pix2Pix can generate images with significantly better quality than other two benchmarks. Among these three compared methods, PSG [18] is the worst. It can be found that PSG does not work sufficiently for recovering OERs as shown in Figure 4. Pix2Pix [29] may underestimate the missing regions, still leaving overexposure in the center of OERs after reconstruction. We quantitatively measure their performances using peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) [40]. PSNR and SSIM are given below each

reconstructed image in Figure 4. The average PSNR and SSIM are also computed on the whole database (over 1600 samples), listed in Table 3. It can be observed that the proposed model outperforms other two benchmarks significantly, achieving up to 5 dB relative to the Pix2Pix [29] and 15 dB relative to the PSG [18].

**4.3. Evaluation on Each Component of Loss Function.** To evaluate contribution of each loss component, we perform the following three testings:

- (1) The adversarial loss and conventional L1 loss on whole image, i.e.,  $\mathcal{L}_{\text{cGAN}}(G, D) + \lambda_1 \mathcal{L}_1(G)$ ;
- (2) The adversarial loss and masked L1 loss, i.e.,  $\mathcal{L}_{\text{cGAN}}(G, D) + \lambda_1 \mathcal{L}_1^m(G)$ ;
- (3) The adversarial loss, masked L1 loss, and edge mask loss/smoothness, i.e.,  $\mathcal{L}_{\text{cGAN}}(G, D) + \lambda_1 \mathcal{L}_1^m(G) + \lambda_2 \mathcal{L}_1^e(G)$ .

Figure 5 illustrates three groups of experiences with different training losses. We can find that despite the outputs of  $\mathcal{L}_{\text{cGAN}} + \lambda_1 \mathcal{L}_1$  have smooth boundary between missing regions and normal regions, it fails to estimate accurate intensity and content of missing regions. Benefiting from the masked loss,  $\mathcal{L}_{\text{cGAN}} + \lambda_1 \mathcal{L}_1^m$  can obtain better intensity and content estimation for missing regions; however, it yields artificial edges as illustrated in the first and third images in Figure 5. The network trained with the mixed  $\mathcal{L}_{\text{cGAN}} + \lambda_1 \mathcal{L}_1^m + \lambda_2 \mathcal{L}_1^e$  losses outperforms the other two, successfully addressing all of the three criteria aforementioned in subsection 3.1.

Another improvement of the proposed model lies in the replacement of Convolution-BatchNorm/ReLU [33] in Pix2Pix [29] by Convolution-SwitchNorm-LReLU/ReLU [32], which is witnessed by the second and third rows of Table 3. In Table 3, the second row lists the average PSNR and SSIM of Pix2Pix with Convolution-BatchNorm-ReLU [33], while the third row is for Pix2Pix with Convolution-SwitchNorm-LReLU/ReLU [32]. It can be found that the third row is obviously better than the second row with

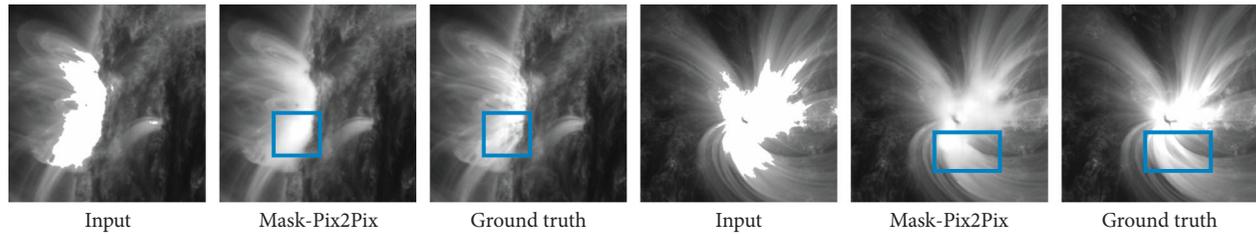


FIGURE 7: Fail examples of our model.

respect to both PSNR and SSIM, which proves the effectiveness of Convolution-SwitchNorm-LReLU/ReLU.

**4.4. Discussions and Testing on Real Solar Image.** In a real AIA image, overexposure region is related to two effects: primary saturation (PS) and secondary saturation (SS) [13]. PS refers to the fact that for intense incoming flux, CCD pixels lose their ability to accommodate additional charge. SS names the fact that PS causes charge to spill into their neighbors. Furthermore, overexposure often accompanies with diffraction, creating a star-like pattern in the images. However, there are no ground-truth images paired with the overexposure ones. Therefore, we train the mask-Pix2Pix model on synthesized database, where the saturated images are cropped from the LSDO.

We further test our pretrained mask-Pix2Pix model on the flare image of Sep. 10, 2017 X8.2. The results are illustrated in Figure 6. The test result shows that the pretrained model recovers the intensity to some extent while fails to recover the fine structure of flare.

## 5. Conclusions

We propose a mask-Pix2Pix network with a hybrid loss function to optimize the recovery of OERs in solar images. The proposed model can robustly handle all kinds of OERs regarding shape, size, and location in an image. In addition, the evaluations on each component of loss function indicate that the two incremental components beyond conventional cGAN have important contributions to the proposed model. Moreover, the investigation on Convolution-SwitchNorm-LReLU/ReLU also demonstrates that better performance can be achieved comparing to conventional Convolution-BatchNorm-ReLU. However, one limitation of our model is that it may fail on highly textured areas, and large OERs, as illustrated in Figure 7, which will be further investigated in our future work.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request. And in the future, the data used to support the findings of this study will be published online.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] K. Shibata, Y. Ishido, L. W. Acton et al., "Observations of X-ray jets with the yohkoh soft X-ray telescope," *Publications of the Astronomical Society of Japan*, vol. 44, pp. L173–L179, 1992.
- [2] F. Clette, J. P. Delaboudiniere, K. P. Dere, and P. Cugnon, "EIT: The Extreme Ultraviolet Imaging Telescope," in *Proceedings of CESRA workshop*, vol. 251, Springer, Postdam, Germany, 1995.
- [3] J.-P. Delaboudiniere, G. Artzner, J. Brunaud et al., "Eit: extreme-ultraviolet imaging telescope for the soho mission," in *The SOHO Mission*, pp. 291–312, Springer, Berlin, Germany, 1995.
- [4] V. Domingo, B. Fleck, and A. I. Poland, "The soho mission: an overview," *The SOHO Mission*, vol. 162, no. 1-2, pp. 1–37, 1995.
- [5] B. N. Handy, L. W. Acton, C. C. Kankelborg et al., "The transition region and coronal explorer," *Solar Physics*, vol. 187, no. 2, pp. 229–260, 1999.
- [6] R. A. Howard, J. D. Moses, A. Vourlidas et al., "Sun earth connection coronal and heliospheric investigation (secchi)," *Space Science Reviews*, vol. 136, no. 1–4, pp. 67–115, 2008.
- [7] M. L. Kaiser, T. A. Kucera, J. M. Davila, O. C. S. Cyr, M. Guhathakurta, and E. Christian, "The stereo mission: an introduction," *The STEREO Mission*, vol. 136, no. 1–4, pp. 5–16, 2008.
- [8] T. Kosugi, K. Matsuzaki, T. Sakao et al., "The hinode (solar-b) mission: an overview," in *The Hinode Mission*, pp. 5–19, Springer, Berlin, Germany, 2007.
- [9] J. R. Lemen, D. J. Akin, P. F. Boerner et al., "The atmospheric imaging assembly (aia) on the solar dynamics observatory (sdo)," in *The Solar Dynamics Observatory*, pp. 17–40, Springer, Berlin, Germany, 2011.
- [10] W. D. Pesnell, B. J. Thompson, and P. C. Chamberlin, "The solar dynamics observatory (sdo)," in *The Solar Dynamics Observatory*, pp. 3–15, Springer, Berlin, Germany, 2011.
- [11] P. Boerner, C. Edwards, J. Lemen et al., "Initial calibration of the atmospheric imaging assembly (aia) on the solar dynamics observatory (sdo)," in *The Solar Dynamics Observatory*, pp. 41–66, Springer, Berlin, Germany, 2011.
- [12] P. H. Scherrer, J. Schou, R. Bush et al., "The helioseismic and magnetic imager (hmi) investigation for the solar dynamics observatory (sdo)," *Solar Physics*, vol. 275, no. 1-2, pp. 207–227, 2012.
- [13] S. Guastavino, M. Piana, A. M. Massone, R. Schwartz, and F. Benvenuto, "Desaturating euv observations of solar flaring storms," 2019, <http://arxiv.org/abs/1904.04211>.
- [14] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1200–1211, 2001.

- [15] S. Esedoglu and J. Shen, "Digital inpainting based on the mumford-shah-euler image model," *European Journal of Applied Mathematics*, vol. 13, no. 4, pp. 353–370, 2002.
- [16] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, "Patchmatch: a randomized correspondence algorithm for structural image editing," *ACM Transactions on Graphics (ToG)*, vol. 28, no. 3, p. 24, 2009.
- [17] S. Darabi, E. Shechtman, C. Barnes, D. B. Goldman, and P. Sen, "Image melding: combining inconsistent images using patch-based synthesis," *ACM Transactions on Graphics*, vol. 31, no. 4, pp. 82–91, 2012.
- [18] J.-B. Huang, S. B. Kang, N. Ahuja, and J. Kopf, "Image completion using planar structure guidance," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 4, pp. 1–10, 2014.
- [19] K. He and J. Sun, "Image completion approaches using the statistics of similar patches," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 12, pp. 2423–2435, 2014.
- [20] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2015, <http://arxiv.org/abs/1511.07122>.
- [21] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: feature learning by inpainting," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2536–2544, Las Vegas, NV, USA, June 2016.
- [22] C. Yang, X. Lu, Z. Lin, E. Shechtman, O. Wang, and H. Li, "High-resolution image inpainting using multi-scale neural patch synthesis," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6721–6729, Honolulu, HI, USA, July 2017.
- [23] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Globally and locally consistent image completion," *ACM Transactions on Graphics (ToG)*, vol. 36, no. 4, pp. 1–14, 2017.
- [24] Y. Song, C. Yang, Z. Lin et al., "Contextual-based image inpainting: infer, match, and translate," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 3–19, Munich, Germany, September 2018.
- [25] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro, "Image inpainting for irregular holes using partial convolutions," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 85–100, Munich, Germany, September 2018.
- [26] Z. Yan, X. Li, M. Li, W. Zuo, and S. Shan, "Shift-net: image inpainting via deep feature rearrangement," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 1–17, Munich, Germany, September 2018.
- [27] Q. Xiao, G. Li, and Q. Chen, "Deep inception generative network for cognitive image inpainting," 2018, <http://arxiv.org/abs/1812.01458>.
- [28] K. Nazeri, E. Ng, T. Joseph, F. Qureshi, and M. Ebrahimi, "Edgeconnect: generative image inpainting with adversarial edge learning," 2019, <http://arxiv.org/abs/1901.00212>.
- [29] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1125–1134, 2017.
- [30] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, <http://arxiv.org/abs/1411.1784>.
- [31] O. Ronneberger, P. Fischer, T. Brox, and U-net, "Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer, Berlin, Germany, 2015.
- [32] P. Luo, J. Ren, and Z. Peng, "Differentiable learning-to-normalize via switchable normalization," 2018, <http://arxiv.org/abs/1806.10779>.
- [33] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," 2015, <http://arxiv.org/abs/1502.03167>.
- [34] J. Lei Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," 2016, <http://arxiv.org/abs/1607.06450>.
- [35] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: the missing ingredient for fast stylization," 2016, <http://arxiv.org/abs/1607.08022>.
- [36] A. Kucuk, J. M. Banda, and R. A. Angryk, "A large-scale solar dynamics observatory image dataset for computer vision applications," *Scientific Data*, vol. 4, no. 1, article 170096, 2017.
- [37] P. Rez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Transactions on Graphics*, vol. 22, no. 3, pp. 313–318, 2003.
- [38] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," <http://arxiv.org/abs/1409.1556>.
- [39] D. P. Kingma, J. Ba, and Adam, "A method for stochastic optimization, international conference on learning representations," 2014, <https://arxiv.org/abs/1412.6980>.
- [40] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

## Research Article

# Intelligent Recognition of Time Stamp Characters in Solar Scanned Images from Film

Jiafeng Zhang <sup>1</sup>, Guangzhong Lin,<sup>1</sup> Shuguang Zeng <sup>1</sup>, Sheng Zheng,<sup>1</sup> Xiao Yang,<sup>2</sup>  
Ganghua Lin,<sup>2</sup> Xiangyun Zeng,<sup>1</sup> and Haimin Wang <sup>3</sup>

<sup>1</sup>College of Science, China Three Gorges University, Yichang 443002, China

<sup>2</sup>Key Laboratory of Solar Activity, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China

<sup>3</sup>Institute for Space Weather Sciences, New Jersey Institute of Technology, 323 Martin Luther King Boulevard, Newark, NJ 07102-1982, USA

Correspondence should be addressed to Shuguang Zeng; zengshuguang19@163.com

Received 11 April 2019; Revised 4 July 2019; Accepted 18 July 2019; Published 28 August 2019

Academic Editor: Geza Kovacs

Copyright © 2019 Jiafeng Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Prior to the availability of digital cameras, the solar observational images are typically recorded on films, and the information such as date and time were stamped in the same frames on film. It is significant to extract the time stamp information on the film so that the researchers can efficiently use the image data. This paper introduces an intelligent method for extracting time stamp information, namely, the convolutional neural network (CNN), which is an algorithm in deep learning of multilayer neural network structures and can identify time stamp character in the scanned solar images. We carry out the time stamp decoding for the digitized data from the National Solar Observatory from 1963 to 2003. The experimental results show that the method is accurate and quick for this application. We finish the time stamp information extraction for more than 7 million images with the accuracy of 98%.

## 1. Introduction

The chromosphere is a layer of atmosphere between the photosphere and the corona. The chromospheric magnetic field structure is high dynamic, and the most intensive activities are solar flares. In order to study the solar flares and other solar activities, it is necessary to accumulate the flare observations in the chromosphere for many years. Therefore, a number of solar telescopes have been established around the world, for example, Solar Magnetic Field Telescope (SMFT) in Huairou [1], China, and McMath-Pierce Solar Telescope in Arizona [2], USA. Prior to the availability of modern digital cameras, the main medium for recording solar chromosphere data was film. In order to use rich historical data, many projects are involved to digitize historical astronomical data, and new research results are obtained from old data such as observation of a Moreton wave and wave-filament interactions associated with the renowned X9 flare on 1990 May 24 [3], circular ribbon flares,

and homologous jets [4]. Because of the huge amount of data, the time stamps of many digitized chromospheric images are still in the form that cannot be read directly by the computer, which has produced obstacles to further research. The digitization of time stamps makes the data to be more efficiently analyzed. Therefore, time stamp decoding is a significant problem that we intend to solve.

From 1963 to 2003, full-disk Halpha images are recorded in 35 mm films with 1 minute or even shorter cadence at National Solar Observatory (NSO) of the US. More than 8 million pictures have been recorded and then digitized by the New Jersey Institute of Technology (NJIT), which covers hundreds of solar flares and other activities. It will create a valuable data archive of solar eruptions, which is a huge advance in solar astronomy. However, the data is useless before the decoding of time stamp. An example of a chromospheric image is shown in Figure 1. The image records some information such as the year, month, day, hour, minute, second, and film number, besides full-disk

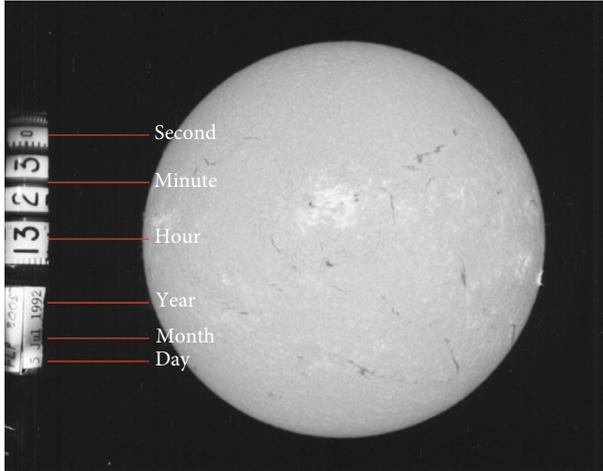


FIGURE 1: Full-disk chromosphere image obtained in Halpha by NSO. The time stamp is on the left side of the image, and the full-disk chromosphere image in the middle.

solar image. The time/date when the picture was taken is what we need to extract. As the amount of data is very large, automatically identifying the characters of the time stamp is the key to the efficient usage of the data. In order to solve the problem of character recognition, many methods have been proposed, such as support vector machine algorithm [5], deep learning algorithm [6–8], and so on.

Recently, the convolutional neural network [9, 10] (CNN) is a popular deep learning algorithm with high accuracy in classification. It has been widely used in face recognition [11], image classification [9], speech recognition [12], character recognition [13, 14], etc. Zheng et al. [13] applied it for character recognition in the sunspot drawings of Yunnan Observatory, with the accuracy of 98.5%. Goodfellow et al. [14] has applied CNN to the Street View House Numbers (SVHNs) dataset with the accuracy of 96%. We adopt the CNN for character recognition because of the high accuracy. The selection effect of samples is the key to the recognition accuracy of the CNN. However, the characters in the time stamp are specific and not included in any digital sample database. We need to create a sample database for them as a training set. In addition, many images are ambiguous, and there is still a big hindrance to solving character segmentation and recognition.

In this paper, we present an intelligent recognition method for automatic segmentation and recognition of characters based on CNN. The paper is organized as follows. Section 2 is an introduction to CNN algorithm. In Section 3, we apply the CNN algorithm to time stamp recognition. Section 4 demonstrates the recognition result of this method for the time stamp. Finally, we give a conclusion in Section 5.

## 2. Convolutional Neural Network

The CNN [9, 10, 15] includes input layer, convolutional layers, pooling layers, fully connected layers, and output layer. A typical structure is shown in Figure 2. Feature vectors in the outer layer are extracted from data in the input

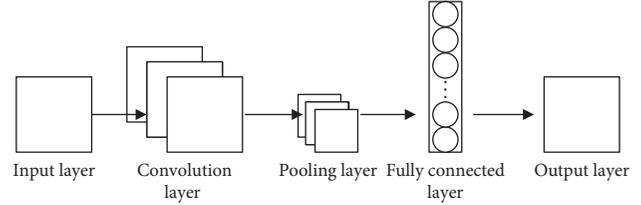


FIGURE 2: Convolutional neural network structure, with input layer, convolution layer, pooling layer, fully connected layer, and output layer.

layer by the convolutional layer, the pooling layer, and the fully connected layer and then used in classifying the input data by logistic regression.

Multiple convolutional layers, pooling layers, and fully connected layers are possible in the CNN. The convolution layer detects the characteristics of the input layer to the maximum extent by randomly generating sufficient convolution kernels. A large number of feature maps are generated after passing through the convolutional layer. The convolution layer is usually followed by an activation function which is used to convert features from a linear space into a nonlinear space to achieve the nonlinear classification [16]. ReLu, sigmod, and tanh are commonly applied as the activation functions. In this paper, ReLu is adopted, which can effectively prevent overfitting problems. The pooling layer is a feature filter for the convolutional layer to preserve the main features and to reduce the amount of computation. It is often placed in the middle of two convolutional layers.

The data, which processed by multiple convolution layers and pooling layers, are connected to one or more fully connected layers. In a fully connected layer, each neuron is connected to all neurons in the upper layer to combine the features extracted previously, so the extracted features can be completely preserved and unaffected by the position in the original image. The output value of the output layers is classified by logistic regression. Softmax regression is usually used when dealing with multiclassification problems. The Softmax regression outputs the probability value of the sample for each class and selects the class corresponding to the maximum probability as the recognition result of the sample. In addition, the recognition accuracy of CNN is closely related to the quality and quantity of samples. The richer the training samples, the higher the recognition accuracy.

## 3. Time Stamp Character Recognition Based on CNN

The information of year, month, day, hour, and minute is what we need to extract in the image. Figures 3 and 4 show chromospheric pictures with two types of the time stamp. The time stamp in Figure 3 is black on white, while Figure 4 is white on black. The time stamps are uneven, the format and color of the characters are inconsistent, the YMD (year, month, and day) characters are small, and the characters in many pictures are illegible and difficult to recognize.

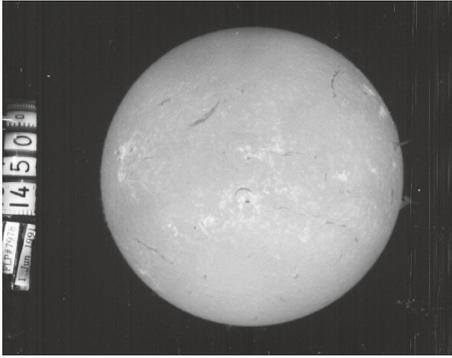


FIGURE 3: Black character image.



FIGURE 4: White character image.

However, the date information is continuous, and there are many images on the same date. So, we only need to get the date of the first picture every day without intelligent recognition. That part was achieved manually. The CNN is used in identifying the HM (hour and minute) characters. The flow chart of the CNN algorithm for recognizing time stamp characters is shown in Figure 5. It consists of two independent parts: one for character segmentation (Section 3.1) and the other for character recognition by CNN (Section 3.2).

The left part of the flow chart introduces image segmentation, and the right part introduces character recognition. The input image is processed by white characters by default. If no character areas can be extracted, the process returns to the binarization. Reverse the color of white and black in the binary image; The CNN will be retrained when it has a low recognition rate for the test samples.

**3.1. Characters Segmentation.** The size of the original image is  $1600 \times 2048$ , as shown in Figure 3. The time stamp is on the left or right side of the picture and the character format is different. Characters are divided into two categories, one is black and the other is white, which need to be dealt with separately. The character segmentation steps are as follows.

*Step 1.* Remove the part of the solar disk from the picture and obtain the left and right sides of the picture.

*Step 2.* Get the picture with a time stamp based on the intensity variance across the picture, and rotate the picture to adjust the direction of the characters (Figure 6(a)).

*Step 3.* Eliminate noise in pictures with top hat operation (Figure 6(b)).

*Step 4.* Binarize the picture by the Sauvola algorithm [17].

*Step 5.* Reserve connectivity domain of which area is in (500, 1000).

*Step 6.* Extract character regions using stroke width transform algorithm [18,19] (Figure 6(e)).

*Step 7.* If there are no character regions, return to step 4. Reverse the color of white and black in the binary image which is obtained in step 4 (Figure 6(c)) to get white

characters, to allow black characters in the original image to be extracted. This ensures data consistency so that the next steps are as identical as possible. After step 5 as shown in Figure 6(d). If there are still no character regions after the image is reversed, it means that there are no characters in the current picture. Because there are only two forms of time stamp and few a part of the images that do not contain time stamps, the time stamp characters cannot be extracted from these images during the above process.

*Step 8.* Extract the corresponding region from the original image according to the binary image, and resize each of the characters to  $28 \times 28$  (Figure 6(f)).

**3.2. Characters Recognition.** The CNN model for time stamp character recognition consists of two convolutional layers, two pooling layers, and a fully connected layer (Figure 7). In the first convolutional layer Con\_1, 6 different convolutional kernels of size  $5 \times 5$  are used to take convolution operation on character pictures with the size of  $28 \times 28$ . After Con\_1, the original character picture becomes a  $24 \times 24 \times 6$  feature map. The first pooling layer Pool\_1 filters the feature map using maximum pooling function with the sliding window of  $2 \times 2$ . Then, it becomes a feature map of size  $12 \times 12 \times 6$ . The convolutional layer Con\_2 contains 10 kernels of  $5 \times 5$ . The pooling layer Pool\_2 does the same as Pool\_1. These feature maps are taken as the inputs into the fully connected layer (F) to obtain the feature vector. Finally, the vector is classified by the softmax function to obtain the recognition result.

The training steps of the CNN in this paper are divided into the following three steps.

*Step 1.* Add labels to the single-character images as samples for training the network.

*Step 2.* The character image is used as the  $X$  vector for the input layer, and the label of the image is used as the  $Y$  vector.

*Step 3.* The network is trained by forwarding propagation and back propagation algorithms, and its coefficients are updated by loop iteration. A network structure with higher recognition accuracy is obtained in the end.

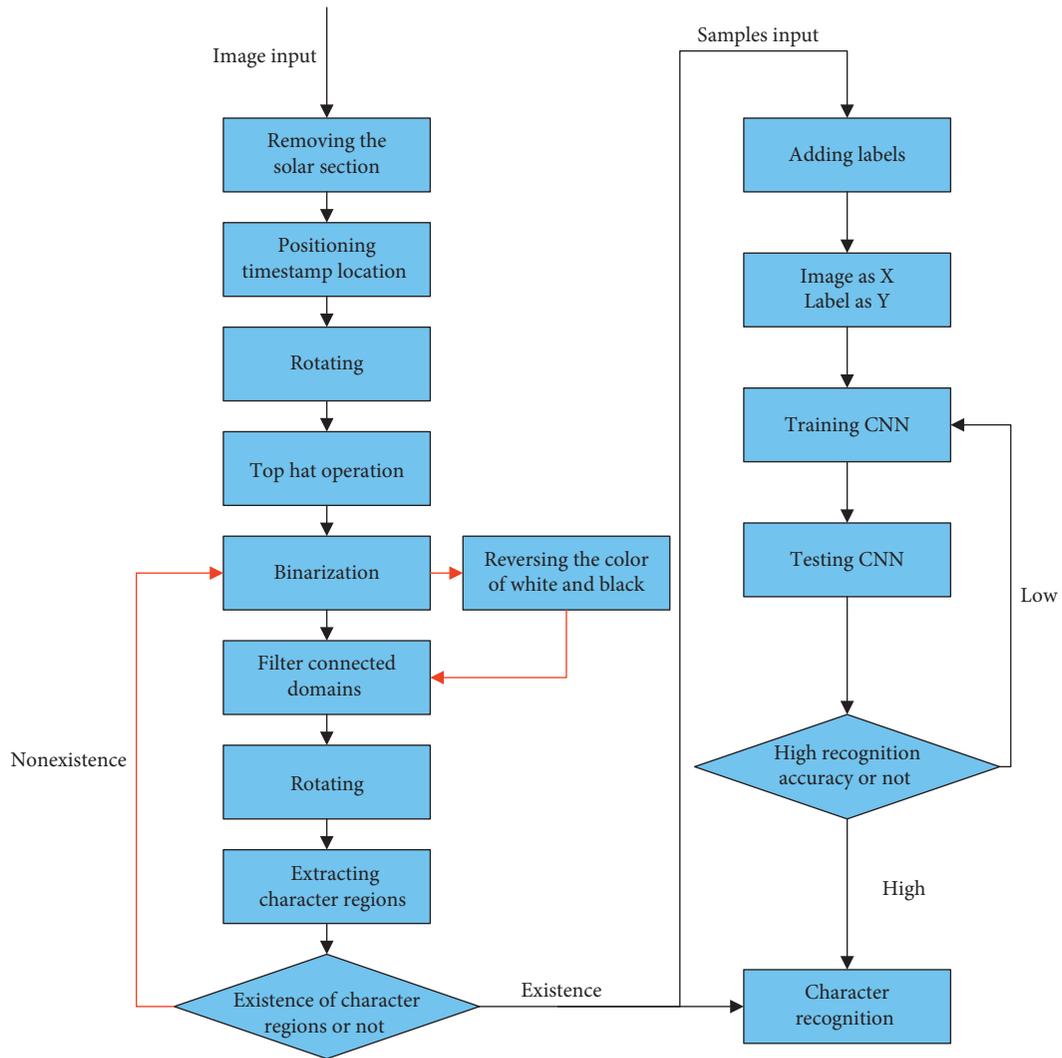


FIGURE 5: Algorithm flow chart.

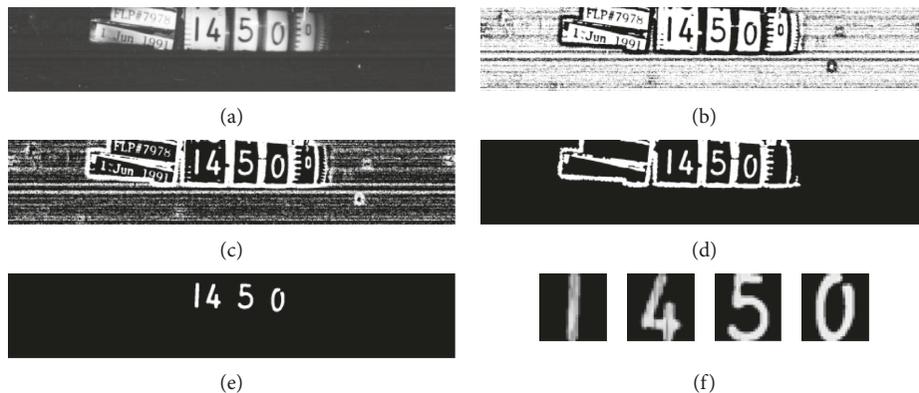


FIGURE 6: Extract characters from Figure 3: (a) position the time stamp and rotate, (b) after top hat operation and binarization, (c) black and white inversion of binary image, (d) extract related connectivity domain, (e) extract character regions, and (f) characters extraction result.

To train the CNN, we select 100,000 single-character images of size  $28 \times 28$ , which were cut from the original images with white characters, as training samples, 10,000

samples per character. These characters are recognizable by humans and labeled manually. There is no need to deal with time stamps unrecognizable by humans, because it is

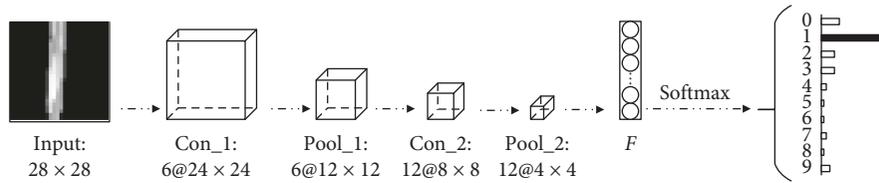


FIGURE 7: Convolution neural network structure of character recognition, included a input layer, two convolution layers, two pooling layers, and a fully connected layer.

impossible to verify the recognition correctness. 9000 images are randomly selected as samples to train the network. The remaining samples are used as testing samples to test the recognition accuracy of the network. The test results are shown in Table 1. From the table, the recognition accuracy of each character is over 98%, and it takes only about 6 seconds to recognize 1000 pictures.

At present, the commonly used methods for character recognition are Optical Character Recognition (OCR) [20] and character recognition based on deep neural network. It is well known that OCR recognizes standard characters effectively. So we did an experiment based on open recognition engine TESSERACT [21]. We train it in the same way as CNN, and the same way to test it. The test results are shown in Table 2 that the highest recognition accuracy is 96.8% and the lowest is 93.2% and the lowest time cost of testing 1000 samples is 8.23 seconds. Convolutional neural networks, on the contrary, have higher recognition accuracy and lower time cost. The reason for the relatively low recognition accuracy of OCR is that characters extracted from time stamps are affected by some interference, such as illumination interference, background noise interference, as shown in Figure 8. It is hard for OCR to handle these situations. So it can be concluded from the comparative experiments that CNN has better robustness, stronger antijamming, and lower time consumption than OCR.

**3.3. Date Check.** After the hour and minute in the time stamp are identified, another important step is to complete the information of the date (year, month, and day). Since the date of the photo may not be continuous and cannot be filled in automatically by the program, it is necessary to confirm the date manually. Although the dates are not continuous, they are all in order, and the volume number, which is recorded in the folder name, of the film helps to determine the range of date. In addition, the photographing time is mostly continuous and the 24-hour timekeeping method is used; it is easy to judge whether the date has changed. For example, if time information of the first picture is “2359” and that of the second picture “000,” the date information of the second picture can be added one day based on the first picture. So for images over a period of time, it only needs to know the observation date of the first picture. However, some dates are not continuous, so a manual check is required. So we adopt a user graphical interface (Figure 9) to assist in the date confirmation. Only the first few pictures of

TABLE 1: Test results of CNN identification.

Character	Total numbers	Recognition errors	Recognition rate	Time cost(s)
0	1000	20	0.980	6.01
1	1000	4	0.996	6.11
2	1000	8	0.992	5.96
3	1000	3	0.997	6.28
4	1000	1	0.999	6.24
5	1000	19	0.981	6.52
6	1000	12	0.988	6.05
7	1000	3	0.997	6.17
8	1000	12	0.988	5.95
9	1000	13	0.987	6.11

TABLE 2: Test results of TESSERACT.

Character	Total numbers	Recognition errors	Recognition accuracy rate	Time cost(s)
0	1000	57	0.943	8.95
1	1000	45	0.955	8.32
2	1000	53	0.947	8.28
3	1000	50	0.950	8.84
4	1000	59	0.941	8.23
5	1000	40	0.960	8.44
6	1000	49	0.951	8.52
7	1000	32	0.968	8.39
8	1000	68	0.932	8.45
9	1000	60	0.940	8.69



FIGURE 8: Character samples affected by illumination and background noise.

a day need to be verified. If the date is incorrect, modify it manually, and the program will automatically update all dates in the subsequent pictures.

Fill in paths of the original image and the record table in the corresponding text box of the program. Click on the “Open” button to open the first image in the folder and its date information is displayed in the corresponding text box. Click on the “Next” or “Last” button to open the next or previous image, respectively. Click on the “Update” button to update the date. The “Next day” button is used to jump directly to the next day. Finally, the updated contents are saved in corresponding files.

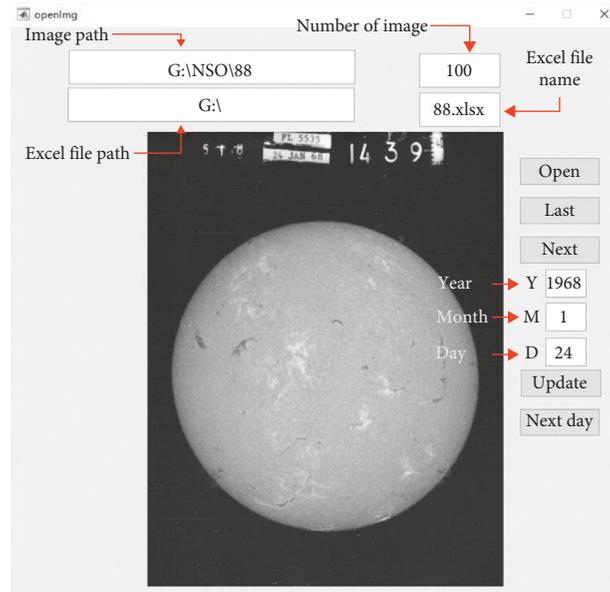


FIGURE 9: Graphic interface of date check.

TABLE 3: Statistical table of recognition results.

	Correct	1 error	2 errors	3 errors	4 errors	Average time cost (s)	Recognition accuracy rate
Numbers	9788	202	10	0	0	0.09	97.9%

TABLE 4: Confusion matrix of character recognition results.

Character	0	1	2	3	4	5	6	7	8	9
0	4558	0	2	108	0	0	16	0	0	1
1	0	11147	0	0	0	0	0	0	0	0
2	0	0	4826	0	0	0	0	0	0	0
3	0	1	85	3890	0	1	0	0	0	0
4	0	0	0	0	4176	2	0	0	0	0
5	0	0	0	0	0	4362	0	0	0	0
6	0	0	0	0	0	0	1299	0	0	0
7	0	0	0	0	0	0	0	1918	0	0
8	0	0	0	0	0	0	0	1	1846	0
9	0	0	0	0	0	0	0	0	1	1760
Total	4558	11148	4913	3998	4176	4365	1315	1919	1847	1761
Number of errors	0	1	87	108	0	3	16	1	1	1
Recognition rate	1	0.99	0.98	0.97	1	0.99	0.99	0.99	0.99	0.99

#### 4. Result and Discussion

To further test the recognition accuracy of the network under actual conditions, we randomly selected 10,000 original images for testing. Table 3 shows the accuracy of the testing results recognized by CNN, which is confirmed manually. Misrecognizing 1 character occurs 202 times, misrecognizing 2 characters occurs 10 times, and no situation occurs for misrecognizing more than 3 characters simultaneously. The recognition accuracy rate is 97.9%, and the average time taken for each picture is 0.09 seconds. The statistics of recognition results for each character are shown in Table 4.

Table 4 shows that the recognition accuracy of the character "0" is 100%, that of "1," "5," and "7" is greater than 99.9%, and that of other characters is above 97.3%. The average recognition accuracy of all the characters is 99.5%. However, the recognition error rates of the characters "2," "3," and "6" are higher, mainly due to these characters being affected by the light, as shown in Figure 10. When they are affected by illumination, they are easily destroyed by the local binary algorithm leading to structural breaks. The character fragments are considered to be noise in the next step of the algorithm because of their small area, which will affect the recognition results (e.g., Figures 10(b) and 10(d)). However, these images affected by lighting only account for a small part

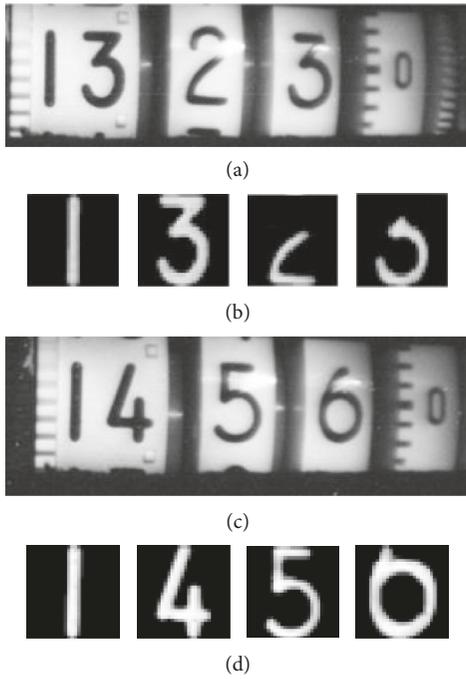


FIGURE 10: Original image (a, c) and segmentation result (b, d). As shown in figure (b, d), the third character and the fourth character are not completely split in (b), the fourth character is not completely split in (d), resulting in incorrect recognition results.

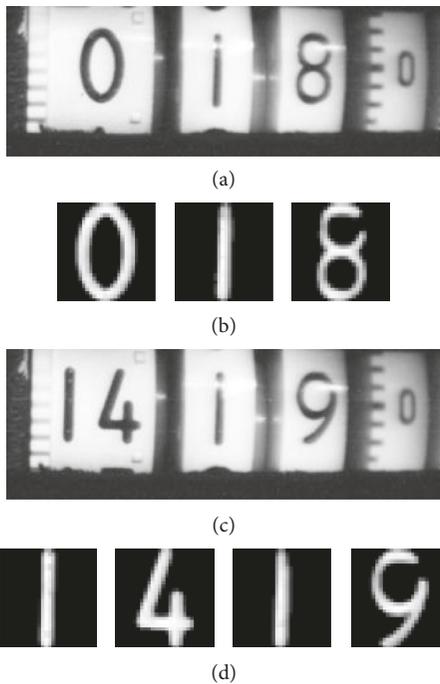


FIGURE 11: Original image (a, c) and segmentation result (b, d), respectively. As shown in figure (b, d), the recognition result is correct, even though the fourth character is partially split.

of the whole samples, as shown in Table 4, so they contribute a little to the average recognition accuracy. Besides, the recognition results of some characters are not affected by illumination, such as “8” and “9,” as shown in Figure 11. When

TABLE 5: Annual total amount of images.

Year	Number
1963	124469
1964	474370
1965	539969
1966	559126
1967	683452
1968	489625
1969	584851
1970	46047
1971	0
1972	16624
1973	350606
1974	289724
1975	213868
1976	182321
1977	<b>190870</b>
1978	207012
1979	293451
1980	302258
1981	236533
1982	188228
1983	178873
1984	174902
1985	32324
1986	0
1987	113652
1988	217226
1989	67403
1990	0
1991	109306
1992	185144
1993	121401
1994	86425
1995	98994
1996	46638
1997	82101
1998	76434
1999	37270
2000	34679
2001	68460
2002	43292
2003	12983
Total	<b>7760911</b>

there is lighting interference on the images, their main structures are preserved so that their recognition results are not affected. The defective structures can be identified, which is one of the advantages of CNN.

Although these images affected by lighting only account for a small part, to solve this problem, our further plan is adding some samples affected by lighting to the training set and improving the algorithm of character segmentation.

In total, we get date/time information for more than 7 million pictures of 38 years, as shown in Table 5. The remaining unprocessed images such as those in 1971, 1986, and 1990 are due to time stamps that are beyond human recognition or do not have time stamps, about 10% of the total. It is not necessary to deal with these pictures because it is impossible to verify whether they are recognized correctly or not. The number of pictures per year is also shown in a bar

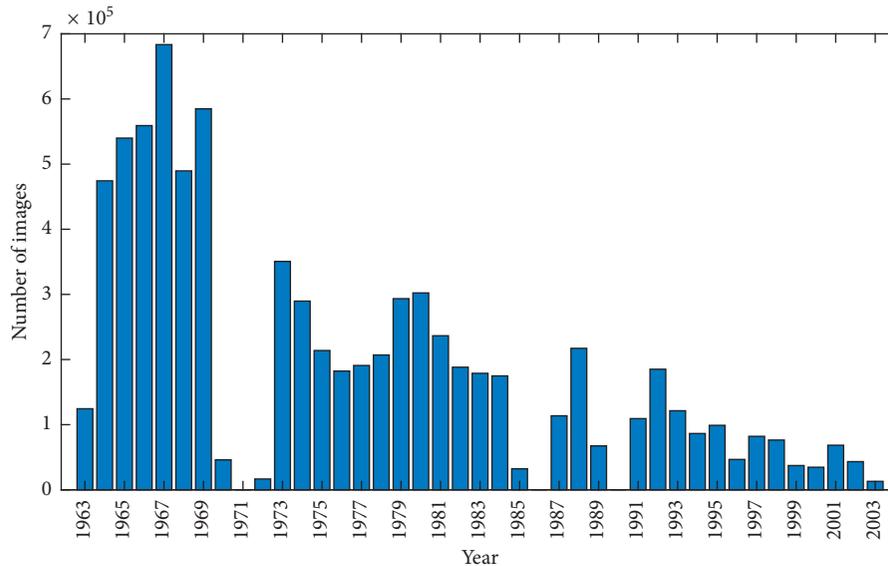


FIGURE 12: Number of pictures per year.

chart as shown in Figure 12. The number of pictures rose slowly from 1963 to 1967, peaking in 1967 with about 700 thousand pictures. After 1967, the number of pictures declined dramatically. In 2003, there were about 13,000 pictures.

## 5. Conclusion

In this paper, we describe an intelligent algorithm to extract the time stamp from traditional films based on CNN. The experimental results show that the method has a good result and meets the speed and quality requirements for identification. It also has strong portability in solving the same type of problems in similar applications.

Finally, we get date/time information for more than 7 million pictures which are recorded by NSO of the US. This greatly reduces the amount of manual work, so that this batch of data can be effectively utilized by researchers as soon as possible. The method proposed in this paper can be applied to character recognition in other historical image, such as handwritten character recognition in sunspot drawing.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request. And in the future, the data used to support the findings of this study will be published online.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work is supported in part by the National Natural Science Foundation of China under Grants U1731124,

U1531247, 11427803, 11427901, and 11873062, the 13th Five-year Informatization Plan of Chinese Academy of Sciences under Grant XXH13505-04, and the Beijing Municipal Science and Technology Project under Grant Z181100002918004. Haimin Wang acknowledges the support of US NSF under grant AGS-1620875. The authors are grateful to the National Solar Observatory for providing the original film data.

## References

- [1] G. Ai and Y. Hu, "On principle of solar magnetic field telescope," *Acta Astronomica Sinica*, vol. 2, pp. 91–98, 1986.
- [2] <https://www.noao.edu/outreach/kptour/mcmath.html>.
- [3] R. Liu, C. Liu, Y. Xu, W. Liu, B. Kliem, and H. Wang, "Observation of a Moreton wave and wave-filament interactions associated with the renowned X9 flare on 1990 May 24," *The Astrophysical Journal*, vol. 773, no. 2, p. 166, 2013.
- [4] H. Wang and C. Liu, "Circular ribbon flares and homologous jets," *The Astrophysical Journal*, vol. 760, no. 2, p. 101, 2012.
- [5] D. Nasien, H. Haron, and S. S. Yuhaziz, "Support vector machine (SVM) for english handwritten character recognition," in *Proceedings of the Second International Conference on Computer Engineering and Applications IEEE*, vol. 1, pp. 249–252, Camastra, Italy, 2010.
- [6] I. H. Witten, E. Frank, M. A. Hall et al., *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, Burlington, MA, USA, 2016.
- [7] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [8] J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.
- [9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the International Conference on Neural Information Processing Systems*, pp. 1097–1105, Curran Associates Inc., Lake Tahoe, NV, USA, December 2012.
- [10] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," *IEEE*

- Transactions on Neural Networks*, vol. 8, no. 1, pp. 98–113, 1997.
- [11] Y. Sun, X. Wang, and X. Tang, “Deep learning face representation from predicting 10,000 classes,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1891–1898, Las Vegas Valley, NV, USA, June 2014.
  - [12] G. Hinton, L. Deng, D. Yu et al., “Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups,” *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 2012.
  - [13] S. Zheng, X. Zeng, G. Lin et al., “Sunspot drawings handwritten character recognition method based on deep learning,” *New Astronomy*, vol. 45, pp. 54–59, 2016.
  - [14] I. J. Goodfellow, Y. Bulatov, J. Ibarz et al., “Multi-digit number recognition from street view imagery using deep convolutional neural networks,” 2013, <https://arxiv.org/abs/1312.6082>.
  - [15] Y. Kim, *Convolutional Neural Networks for Sentence Classification*, 2014.
  - [16] J. Gu, Z. Wang, J. Kuen et al., “Recent advances in convolutional neural networks,” *Pattern Recognition*, vol. 77, pp. 354–377, 2018.
  - [17] J. Sauvola and M. Pietikäinen, “Adaptive document image binarization,” *Pattern Recognition*, vol. 33, no. 2, pp. 225–236, 2000.
  - [18] B. Epshtein, E. Ofek, and Y. Wexler, “Detecting text in natural scenes with stroke width transform,” in *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2963–2970, San Francisco, CA, USA, June 2010.
  - [19] Y. Li and H. Lu, “Scene text detection via stroke width,” in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR 2012)*, pp. 681–684, IEEE, Tsukuba, Japan, November 2012.
  - [20] K. M. Mohiuddin and J. Mao, “Optical character recognition,” in *Wiley Encyclopedia of Electrical and Electronics Engineering*, Springer, Berlin, Germany, 1999.
  - [21] R. Smith, “An overview of the tesseract OCR engine,” in *Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, IEEE Computer Society, Curitiba, Parana, Brazil, September 2007.

## Research Article

# Full-Disk Solar Flare Forecasting Model Based on Data Mining Method

Rong Li<sup>1</sup> and Yong Du<sup>2</sup>

<sup>1</sup>*School of Information, Beijing Wuzi University, Beijing 101149, China*

<sup>2</sup>*Department of Electrical and Information Engineering, Northeast Agricultural University, Harbin, China*

Correspondence should be addressed to Rong Li; [lirong@bao.ac.cn](mailto:lirong@bao.ac.cn)

Received 11 April 2019; Accepted 18 June 2019; Published 1 August 2019

Guest Editor: Liyun Zhang

Copyright © 2019 Rong Li and Yong Du. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Solar flare is one of the violent solar eruptive phenomena; many solar flare forecasting models are built based on the properties of active regions. However, most of these models only focus on active regions within 30° of solar disk center because of the projection effect. Using cost sensitive decision tree algorithm, we build two solar flare forecasting models from the active regions within 30° of solar disk center and outside 30° of solar disk center, respectively. The performances of these two models are compared and analyzed. Merging these two models into a single one, we obtain a full-disk solar flare forecasting model.

## 1. Introduction

Solar activities are the primary source of space weather. As one of the important solar eruptive phenomena, solar flares associated with the electromagnetic radiation and energetic particles often interfere with geostationary satellites, communication systems, and even power grids [1]. Therefore, solar flare forecasting is a significant topic in space weather forecasting community.

Because the trigger mechanisms of solar flares are unsolved, the current solar flare forecasting only depends on the probabilistic model. The statistical and data mining methods are used to build solar flare forecasting models. Miller (1989) developed an expert system (WOLF) to forecast the probable occurrence of solar flares [2]. McIntosh (1990) summarized the McIntosh classifications of sunspots and built an expert system (Theo) to forecast X-ray flares [3]. Long after this work, the McIntosh classifications are considered as a guide in forecasting solar flares in many space weather prediction centers. Measuring contributions of the McIntosh classifications for solar flare forecasting, Bornmann and Shaw (1994) built a solar flare forecasting model using multiple liner regression analysis [4]. Wheatland (2004) pointed out that the history of solar flares is also an important indicator for the occurrence of solar flares, so a Bayesian approach was

proposed to forecast solar flares using the previous flaring records [5]. Leka and Barnes (2007) applied discriminant analysis to produce a binary categorization for the flaring and nonflaring regions [6], and this approach was extended to a probabilistic forecast in Barnes et al. (2007) [7].

Data mining methods also have a long history for the application in solar flare forecasting. Bradshaw et al. (1989) trained a three-layer neural network to forecast flares [8]. Wang et al. (2008) built a solar flare forecasting model supported with an artificial neural network based on the solar magnetic field parameters [9]. Li et al. (2007) proposed a data mining method combining the support vector machine and the k-nearest neighbors to train a solar flare forecasting model [10]. Qahwaji and Colak (2007) built a hybrid system that combines a support vector machine and a cascade-correlation neural network for solar flare forecasting [11]. The sequential information of active regions is analyzed in [12–16]. The active longitudes information is used to improve the performance of solar flare forecasting in [17]. At present, deep learning methods have been used to build solar flare forecasting models [18, 19].

Because of the projection effect of solar magnetograms, active regions within 30° of solar disk center, where projection effect can be negligible, are usually selected to extract parameters and furthermore to build the forecasting model.

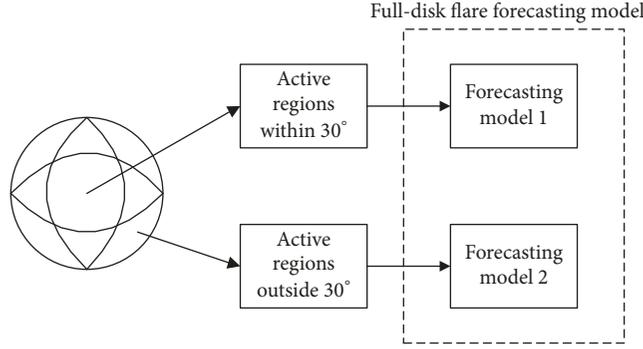


FIGURE 1: Full-disk solar flare forecasting model.

However, active regions which locate outside  $30^\circ$  of solar disk center also produce solar flares. In the present work, we collect the data for active regions outside  $30^\circ$  of solar disk center and their related solar flares and build a solar flare forecasting model from this dataset. Combining it with the solar flare forecasting model trained from active regions within  $30^\circ$  of solar disk center, we obtain a full-disk solar flare forecasting model shown in Figure 1.

The paper is organized as follows. In Section 2, we introduce active region parameters and the related flare catalog. In Section 3, we describe the data mining method. In Section 4, we estimate the performance of the solar flare forecasting model. And finally, in Section 5, we give a brief summary of this work.

## 2. Data

**2.1. Active Region Data.** The Solar Dynamics Observatory (SDO) satellite is launched on 2010 February. The Helioseismic and Magnetic Imager (HMI), which is one of three instruments aboard the SDO, measures the full-disk photospheric vector magnetic field [20]. In 2014, a data product called Space Weather HMI Active Region Patches (SHARP) automatically identifies active regions using the vector magnetic field data when these active regions cross the solar disk [21]. For this study, we use the active region vector magnetic field data generated by the SDO's SHARP data patches from 2011 August to 2012 July. We calculate 4 physical parameters using these 12 month vector magnetic field data, and obtain 2966 samples including 1436 samples within  $30^\circ$  of solar disk center and 1530 samples outside  $30^\circ$  of solar disk center.

The 4 physical parameters are:

- (1) The maximum horizontal gradient of the longitudinal magnetic field: this parameter estimates maximum squeezing among flux systems in an active region.
- (2) The length of neutral lines: the neutral lines separate opposite polarities of the longitudinal magnetic field [22].
- (3) The number of singular points: it is the number of nodes in the network formed by magnetic separatrices [22].

- (4) Sum of photospheric magnetic free energy.

$$\rho_{\text{sum}} = \sum (B^{\text{obs}} - B^{\text{pot}})^2 \quad (1)$$

$\rho_{\text{sum}}$  measures the nonpotentiality of an active region.

**2.2. Flare Data.** According to the peak flux of 1 to 8 angstrom X-rays, solar flare is classified as different class levels shown in Table 1. Within a class level, there is a linear scale from 1 to 9. For example, a C2 flare is twice as powerful as a C1 flare.

Solar flares whose Geostationary Operational Environmental Satellite (GOES) X-ray flux peak magnitude is larger than the C1.0 level are considered in the present work. Solar flare data is collected from the National Geophysical Data Center GOES X-ray flux flare catalogs. An active region is considered as a flaring sample, when this region produces a flare whose level is larger than C1.0 within 48 hours after the observation of this active region. Otherwise, an active region is considered as a nonflaring sample. As such, there are 74 flaring samples and 1362 nonflaring samples for active regions within  $30^\circ$  of solar disk center. And there are 101 flaring samples and 1429 nonflaring samples for active regions outside  $30^\circ$  of solar disk center.

## 3. Method

**3.1. Basic Algorithm.** A decision tree is a flowchart-like model that shows the various outcomes from a series of decisions. It can be used for research analysis or for building forecasting model.

Decision trees have three main parts: a root node, leaf nodes, and branches. The root node is the starting point, root contains questions or criteria to be answered, and leaf nodes stand for the decision of the model. Branches are arrows connecting nodes, showing the information flow between the nodes.

The decision tree algorithm is used to build the solar flare forecasting model. This means that the forecasting model will be represented by a tree-like structure shown in Figure 2 [23]. The decision tree consists of testing nodes, leaf nodes, and branches. A sample is classified from the root node. The specified parameter of this node is calculated and the sample is moved down along the corresponding branch and

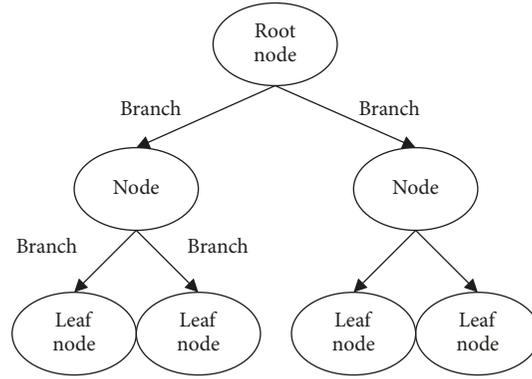


FIGURE 2: Structure of decision tree.

TABLE 1: Classifications of solar X-ray flares.

Class level	peak flux of 1 to 8 angstrom X-rays (Watts/square metre)
A	$< 10^{-7}$
B	$10^{-7} - 10^{-6}$
C	$10^{-6} - 10^{-5}$
M	$10^{-5} - 10^{-4}$
X	$> 10^{-4}$

finally goes to the leaf node where the classification result is determined.

The decision tree is constructed from the training set recursively. In each step, the best parameter is selected to generate the test node and the corresponding branches. The parameter is evaluated by information gain ratio

$$GR(D, F) = \frac{IG(D, F)}{H(F)} \quad (2)$$

where  $D$  stands for the decision of the model,  $F$  stands for the feature of the model,  $IG(D, F) = H(D) - H(D|F)$  is the information gain (IG), and  $H$  stands for the entropy which is used to measure the uncertainty of a system.

The training dataset is divided into some subsets according to the value of branches. This process is repeated until the following stop criteria are satisfied: (1) samples in the subset have the same class label or (2) all possible tests have the same class distribution [24]. When the stop criteria are satisfied, the leaf node is generated. The class label of the samples in the leaf node is the same as that of the majority of samples in this leaf node.

**3.2. Cost Sensitive Modification for the Basic Algorithm.** As shown in Section 2, the ratio between nonflaring samples and the flaring samples is 16. This is called class imbalance problem in data mining community. In order to treat the class imbalance problem, we modified the basic algorithm to the cost sensitive one [25].

In the basic decision tree algorithm, the probability is a basic component to calculate the entropy, information

gain, and information gain ratio. Generally, the probability is estimated by the frequency calculated from the dataset.

$$P(D = d_i) = \frac{|D = d_i|}{|D|} \quad (3)$$

where  $|D|$  is the number of samples in set  $D$  and  $|D = d_i|$  is the number of samples with class label  $d_i$  in set  $D$ .

In the cost sensitive algorithm, there are different costs for different class labels. For example, for a binary classification problem, the cost for class  $d_0$  is  $C_0$ , and the cost for class  $d_1$  is  $C_1$ . Thus, the probability for cost sensitive problem can be estimated as follows.

$$P_{cost}(D = d_0) = \frac{|D = d_0| \times C_0}{|D = d_0| \times C_0 + |D = d_1| \times C_1} \quad (4)$$

$$P_{cost}(D = d_1) = \frac{|D = d_1| \times C_1}{|D = d_0| \times C_0 + |D = d_1| \times C_1} \quad (5)$$

In fact, the usual probability is considered as the cost sensitive probability when the costs  $C_0$  and  $C_1$  are settled to 1. Using the cost sensitive probability, we can calculate the cost sensitive entropy and information gain. Similar to the basic decision tree algorithm, we can build the cost sensitive decision tree model.

## 4. Performance

**4.1. Performance Metrics.** For a binary forecasting model, the results can be summarized in contingency table shown in Table 2. The flaring sample is called positive one, and the nonflaring sample is called negative one. The actual positive sample correctly forecasted as positive one is called true

TABLE 2: Definition of contingency table.

	Forecast positive	Forecast negative
Actual positive	$N_{TP}$	$N_{FN}$
Actual negative	$N_{FP}$	$N_{TN}$

TABLE 3: Contingency table for model 1.

	Forecast positive	Forecast negative
Actual positive	53	21
Actual negative	161	1201

TABLE 4: Contingency table for model 2.

	Forecast positive	Forecast negative
Actual positive	78	23
Actual negative	464	965

TABLE 5: Performances of solar flare forecasting models.

	TP rate	TN rate	HSS
Model 1	71.6%	88.2%	0.316
Model 2	77.2%	67.5%	0.148

positive (TP), the actual positive sample wrongly forecasted as negative one is false negative (FN), the actual negative sample correctly forecasted as negative one is true negative (TN), and the true negative sample wrongly forecasted as positive one is false positive (FP).

Using the contingency table, 3 performance metrics are defined to compare the performance of the forecasting model. The TP rate and TN rate are defined to evaluate the accuracy of flaring samples and nonflaring samples, respectively.

$$TPrate = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (6)$$

$$TNrate = \frac{N_{TN}}{N_{TN} + N_{FP}} \quad (7)$$

Heidke skill score (HSS) is used to evaluate the increase in forecasting power over that of random forecast:

$$HSS = \frac{PC - E}{1 - E} \quad (8)$$

where  $PC = (N_{TP} + N_{TN}) / (N_{TP} + N_{TN} + N_{FN} + N_{FP})$  and

$$E = \frac{(N_{TP} + N_{FN})(N_{TP} + N_{FP})}{(N_{TP} + N_{TN} + N_{TP} + N_{FP})^2} + \frac{(N_{TN} + N_{FP})(N_{TN} + N_{FN})}{(N_{TP} + N_{TN} + N_{FN} + N_{FP})^2}. \quad (9)$$

**4.2. Results.** There are 2966 samples in the dataset. In order to make good use of this data, leave-one-out cross validation method is used to evaluate the performance of the forecasting model. In this method, all but one of the samples is used as

training set, and only one sample is used as testing set. The process is repeated as many times as the number of samples in the dataset. Leave-one-out cross validation method does not waste data; however, it is computationally expensive.

Cost sensitive decision tree is an efficient algorithm, so we can complete the leave-one-out testing. The cost for flaring samples is 50 times larger than that for nonflaring samples.

In order to simplify the following discussion, solar flare forecasting model learned from samples within the 30° of solar disk center is called model 1. And solar flare forecasting model learned from samples outside the 30° of solar disk center is called model 2. The contingency tables of model 1 and model 2 are shown in Tables 3 and 4. Based on these contingency tables, the performances of the two forecasting models can be compared by the performance metrics shown in Table 5.

From Table 5, we can find that the performance of model 2 is worse than that of model 1, because the physical parameters used in model 2 could be influenced by the projection effect. However, the performance of model 2 is acceptable. Combining model 1 and model 2, we can obtain a full-disk solar flare forecasting model.

At present, little work can provide forecasting results of solar flares in the active region beyond 30 degrees of the solar disk; hence, we choose the forecasting results in the active region within 30 degrees to compare them with the flare forecasting results provided by the convolution neural network [24]. The results are shown in Table 6. We find that the flare forecasting model built by the convolution neural network has a higher TP rate, while our forecasting model has a higher TN rate. Because the proportions of flaring samples and nonflaring samples are different in the testing dataset, the HSS is incomparable.

TABLE 6: Performance comparisons.

Performance index	Decision tree	CNN
TP rate	71.6%	85.0%
TN rate	88.2%	81.0%
HSS	0.316	0.143

## 5. Conclusion

Space Weather HMI Active Region Patches data product automatically identifies the active regions when they cross the solar disk. We classify the active region samples into two groups by their location information. The active region samples located within the  $30^\circ$  of solar disk center are classified into group one, and the rest of samples are classified into group two. The projection effect of the samples in group one can be negligible, but the magnetic parameters extracted from active region in group two could not be too accurate because of the projection effect. Two solar flare forecasting models are built using data mining method from two group samples, respectively. The performances of these two forecasting models are estimated. The performance of the forecasting model learned from samples within the  $30^\circ$  of solar disk center is better than that of the forecasting model learned from other samples, because the parameters extracted from the active regions outside the  $30^\circ$  of solar disk center are not accurate enough, and the uncertainty is introduced to evaluate the nonpotentiality of these active regions. A full-disk solar flare forecasting model is generated by combining the two models together.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The data used herein was made possible by funding to Nwra from NASA/LWS contract NNH09CE72C (Dr. Graham Barnes, PI). This work is supported by the National Natural Science Foundation of China (NSFC) (Grant No. 11303051), Beijing Intelligent Logistics System Collaborative Innovation Center, and Beijing Key Laboratory (No. BZ0211).

## References

- [1] G. Ai, H. Wang, and J. Wang, "What is a solar electromagnetic storm?" *Space Weather Journal*, vol. 10, no. 9, 2012.
- [2] A. Heck and F. Murtagh, "Knowledge-based systems in astronomy," in *Lecture Notes in Physics*, 1989, 329.
- [3] P. S. McIntosh, "The classification of sunspot groups," *Solar Physics*, vol. 125, no. 2, pp. 251–267, 1990.
- [4] P. L. Bornmann and D. Shaw, "Flare rates and the McIntosh active-region classifications," *Solar Physics*, vol. 150, no. 1-2, pp. 127–146, 1994.
- [5] M. S. Wheatland, "A bayesian approach to solar flare prediction iop-2016.pngA publishing partnershipA Bayesian Approach to Solar Flare Prediction," *The Astrophysical Journal*, vol. 609, p. 1134, 2004.
- [6] K. D. Leka and G. Barnes, "Photospheric magnetic field properties of flaring versus flare-quiet active regions. iv. a statistically significant sample," *The Astrophysical Journal*, vol. 656, no. 2, p. 1173, 2007.
- [7] G. Barnes, K. D. Leka, E. A. Schumer, and D. J. Della-Rose, "Probabilistic forecasting of solar flares from vector magnetogram data," *Space Weather*, vol. 5, no. 9, p. S09002, 2007.
- [8] G. Bradshaw, R. Fozzard, and L. Ceci, "A connectionist expert system that actually works," *Adv Neu Inform Proc Sys*, vol. 1, pp. 248–255, 1989.
- [9] H. N. Wang, Y. M. Cui, R. Li, L. Y. Zhang, and H. Han, "Solar flare forecasting model supported with artificial neural network techniques," *Advances in Space Research*, vol. 42, no. 9, p. 1464, 2008.
- [10] R. Li, H. N. Wang, H. He, Y. M. Cui, and Z. L. Du, "Support vector machine combined with k-nearest neighbors for solar flare forecasting," *Chinese Journal of Astronomy and Astrophysics*, vol. 7, no. 3, p. 441, 2007.
- [11] R. Qahwaji and T. Colak, "Automatic short-term solar flare prediction using machine learning and sunspot associations," *Solar Physics*, vol. 241, p. 195, 2007.
- [12] D. Yu, X. Huang, H. Wang, and Y. Cui, "Short-term solar flare prediction using a sequential supervised learning method," *Solar Physics*, vol. 255, no. 1, pp. 91–105, 2009.
- [13] D. Yu, X. Huang, Q. Hu, R. Zhou, H. Wang, and Y. Cui, "Short-term solar flare prediction using multiresolution predictors," *The Astrophysical Journal*, vol. 709, no. 1, p. 321, 2010.
- [14] D. Yu, X. Huang, H. Wang, Y. Cui, Q. Hu, and R. Zhou, "Short-term solar flare level prediction using a bayesian network approach," *The Astrophysical Journal*, vol. 710, no. 1, p. 869, 2010.
- [15] X. Huang, D. Yu, Q. Hu, H. Wang, and Y. Cui, "Short-term solar flare prediction using predictor teams," *Solar Physics*, vol. 263, no. 1-2, pp. 175–184, 2010.
- [16] X. Huang and H. N. Wnag, "Solar flare prediction using highly stressed longitudinal magnetic field parameters," *Research in astronomy and astrophysics*, vol. 13, no. 3, pp. 351–358, 2013.
- [17] X. Huang, L. Zhang, H. Wang, and L. Li, "Improving the performance of solar flare prediction using active longitudes information," *Astronomy & Astrophysics*, vol. 549, article A127, p. 6, 2013.
- [18] X. Huang, H. Wang, L. Xu, J. Liu, R. Li, and X. Dai, "Deep learning based solar flare forecasting model. i. results for line-of-sight magnetograms," *The Astrophysical Journal*, vol. 856, no. 1, p. 7, 2018.
- [19] N. Nishizuka, K. Sugiura, Y. Kubo, M. Den, and M. Ishii, "Deep flare net (defn) model for solar flare prediction," *The Astrophysical Journal*, vol. 858, no. 2, 2018.

- [20] J. Schou, P. H. Scherrer, R. I. Bush et al., “Design and ground calibration of the helioseismic and magnetic imager (hmi) instrument on the solar dynamics observatory (sdo),” *Solar Physics*, vol. 275, no. 1-2, pp. 229–259, 2012.
- [21] M. G. Bobra, X. Sun, J. T. Hoeksema et al., “The helioseismic and magnetic imager (hmi) vector magnetic field pipeline: sharps – space-weather hmi active region patches,” *Solar Physics*, vol. 289, no. 9, pp. 3549–3578, 2014.
- [22] Y. Cui, R. Li, L. Zhang, Y. He, and H. Wang, “Correlation between solar flare productivity and photospheric magnetic field properties,” *Solar Physics*, vol. 237, p. 45, 2006.
- [23] X. Huang, H. N. Wang, and X. H. Dai, “Science china physics,” *Mechanics and Astronomy*, vol. 55, no. 10, pp. 1956–1962.
- [24] J. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, Calif, USA, 1993.
- [25] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann Publishers, San Mateo, Calif, USA, 2005.

## Research Article

# Deep Learning for Automatic Recognition of Magnetic Type in Sunspot Groups

Yuanhui Fang <sup>1,2</sup>, Yanmei Cui <sup>1</sup>, and Xianzhi Ao<sup>1</sup>

<sup>1</sup>National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China

<sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, China

Correspondence should be addressed to Yanmei Cui; [ymcui@nssc.ac.cn](mailto:ymcui@nssc.ac.cn)

Received 11 April 2019; Revised 16 June 2019; Accepted 25 June 2019; Published 1 August 2019

Academic Editor: Geza Kovacs

Copyright © 2019 Yuanhui Fang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sunspots are darker areas on the Sun's photosphere and most of solar eruptions occur in complex sunspot groups. The Mount Wilson classification scheme describes the spatial distribution of magnetic polarities in sunspot groups, which plays an important role in forecasting solar flares. With the rapid accumulation of solar observation data, automatic recognition of magnetic type in sunspot groups is imperative for prompt solar eruption forecast. We present in this study, based on the SDO/HMI SHARP data taken during the time interval 2010-2017, an automatic procedure for the recognition of the predefined magnetic types in sunspot groups utilizing a convolutional neural network (CNN) method. Three different models (A, B, and C) take magnetograms, continuum images, and the two-channel pictures as input, respectively. The results show that CNN has a productive performance in identification of the magnetic types in solar active regions (ARs). The best recognition result emerges when continuum images are used as input data solely, and the total accuracy exceeds 95%, for which the recognition accuracy of Alpha type reaches 98% while the accuracy for Beta type is slightly lower but maintains above 88%.

## 1. Introduction

In 1908, Hale used the principle of the Zeeman effect to calculate the magnetic field strength inside sunspots and found that it is stronger than that of the surrounding area [1]. Because the convection in the magnetized plasma is suppressed by the strong magnetic field, temperature and radiation of sunspots are lower than those of their surroundings. Sunspots appear as dark areas in the white light images and are relatively temporary phenomena that can last for a few days to several weeks or even months. A well-developed sunspot is composed of a dark center and a lighter black part around it. The former is the umbra and the latter is the penumbra. As one kind of noticeable phenomena, sunspots not only represent the overall level of solar activities in a certain period, but also give clues to reveal other solar events.

Solar flares are one of the most violent phenomena occurring in the solar atmosphere. Long-term observations show that most of solar flares appear above sunspot groups. However, the mechanism that causes solar flares is far from fully understood. Numerous literatures have focused on the

relationship between sunspot groups and solar flares [2–9]. The characteristics of sunspot groups are important factors for forecasting solar flares. Practically, many parameters, including magnetic gradient, magnetic shear, sunspot numbers, the Zernike moments of magnetograms, and McIntosh sunspot classifications, have been used [10–13]. In particular, the morphology, that is, magnetic type of sunspot groups, has always been an important perspective in the prediction of solar flares. Atac (1987) [14] found that the sunspots producing large flares are of type Dki or Eki with magnetic class D. McIntosh (1990) [15] suggested a 3-component McIntosh classification to describe the complexity of sunspots, yielding the establishment of an expert system for predicting X-ray solar flares. S. Eren et al. (2017) [16] investigated the sunspot and solar flare data from 1996 to 2014, covering a total of 4262 active regions (ARs). Their results showed that large and complex sunspot groups have the flare-production potential about eight times higher than that of small and simple ARs.

Nowadays, both identification and classification of sunspots are mainly carried out manually by experts, which is a subjective, time-consuming, and labor-intensive process.

On the other hand, the increasing number of space missions has led to a rapid accumulation of solar activity data set. Therefore, it has been on the agenda for quite a while to implement automatic identification procedures for sunspot magnetic types. Although much work has been done on the automatic recognition of sunspots [17–19], there are few on automatic classification of sunspot magnetic types by machine learning method. Sinh Hoa Nguyen et al. (2005) [20] applied hierarchical learning method based rough set theory to the problem of classifying sunspots from satellite images. Trung Thanh Nguyen et al. (2006) [21, 22] used machine learning techniques to classify sunspot groups based on the seven-class Modified Zurich scheme. Colak et al. (2008) [23] adopted image processing algorithms and fully connected neural networks to automatically detect and classify sunspots based on the McIntosh classification system. Mehmood A. Abd et al. (2010) [24] employed Support Vector Machines to achieve automatic classification of sunspot groups on full disk white light solar images. Sreejith Padinhatteeri et al. (2016) [25] have suggested a system called Solar Monitor Active Region Tracker-Delta Finder, which has a good performance on  $\delta$  magnetic type recognition. However, the Mount Wilson classification in sunspot groups is lacking of overall automatic identification procedures.

In this work, based on the Mount Wilson classification system, the magnetic types of ARs are labeled by one of the unipolar group Alpha, the bipolar group Beta, and other complex multipole groups, Beta-x. These three magnetic types are identified automatically by using the convolutional neural network (CNN) method with SDO/HMI data taken during the time interval 2010-2017. A training model with high recognition accuracy is obtained. The paper is organized as follows: the Mount Wilson magnetic classification is briefly described in Section 2; both the data source and the preprocessing method are explained in Section 3; the structure of CNN is illustrated in Section 4 together with the training results of models using different input data; Section 5 concludes the paper.

## 2. Mount Wilson Sunspot Classification Scheme

In 1919, the Mount Wilson Observatory in California developed a classification scheme of sunspots according to the polarity of the corresponding magnetic fields [26]. It considers bipolar sunspot groups as a basic type, and other types are regarded as deformations of the bipolar sunspot group. The category is listed below (<https://www.spaceweatherlive.com>):

- (1) Alpha: a unipolar sunspot group
- (2) Beta: a sunspot group having both positive and negative magnetic polarities, with a simple and distinct division between opposite polarities
- (3) Gamma: a complex type of AR. Unlike the Beta class, the distribution of positive and negative polarities of this type is very irregular

- (4) Beta-Gamma: a bipolar sunspot group with sufficiently complex distribution and no obvious boundaries between opposite polarities
- (5) Delta: a bipolar sunspot group with opposite polarity umbrae sharing one penumbra that spans less than 2 degrees with respect to the solar center
- (6) Beta-Delta: a sunspot group of Beta magnetic classification containing one or more Delta sunspots
- (7) Beta-Gamma-Delta: a sunspot group of Beta-Gamma magnetic classification containing one or more Delta sunspots
- (8) Gamma-Delta: a sunspot group of Gamma magnetic classification containing one or more Delta sunspots.

From May, 2010, to May, 2017, a total number of 11306 magnetic type classification records covering 1592 ARs are contained in the Solar Region Summary (SRS) text file, which can be downloaded from the NOAA/SWPC website <https://www.swpc.noaa.gov/>. As shown in Table 1, Alpha magnetic type appears 3576 times, accounting for one-third of the total records, and Beta magnetic type appears 6127 times, accounting for 54%. Complex magnetic types: Beta-Gamma, Beta-Gamma-Delta, and Beta-Delta, appear 1143, 385, and 75 times, respectively, accounting for a small portion of the total records. In addition, the Gamma, Delta, and Gamma-Delta magnetic types have not been reported in SRS files during this period, therefore they are not listed in Table 1.

Statistically, large flares are more likely to occur in ARs with complex magnetic types, while the ARs of Alpha type have a lower probability of flare eruption [5]. In this study, the sunspot magnetic type falls into three categories, the unipolar group Alpha, the bipolar group Beta, and other complex multipole groups, called Beta-x.

## 3. Data Preprocessing

The data used in this paper are observed by SDO/HMI [27, 28]. The 720s SHARP data (*hmi.sharp\_720s - Space-Weather HMI Active Region Patch*) [29] are selected, which include magnetograms and continuum images with the time cadence of 12 minutes. The SHARP data provide active region maps in patches that encompass automatically tracked magnetic concentrations for their entire lifetime. All of the data files we have used are downloaded from the website <http://jsoc.stanford.edu/> and are in Flexible Image Transport System (FITS) format. The data selection satisfies the following criteria:

- (i) The time range is from May, 2010, to May, 2017.
- (ii) In order to guarantee enough variations between the closest AR images, the AR data are taken every 96 minutes.
- (iii) The location range of the data is within the  $\pm 75$  heliolongitude degrees from the solar disk center to reduce the influence of projection.
- (iv) Only when one SHARP number corresponds to one NOAA AR at the same time, the SHARP data is selected.

TABLE 1: Distribution of magnetic types in 1592 ARs recorded in SRS files.

Type	$\alpha$	$\beta$	$\beta\text{-}\gamma$	$\beta\text{-}\gamma\text{-}\delta$	$\beta\text{-}\delta$	Total
Number	3576	6127	1143	385	75	11306
Ratio	31.6%	54.2%	10.1%	3.4%	0.6%	100%

TABLE 2: Distribution of magnetic types of SDO/HMI magnetograms and continuum images.

	Alpha	Beta	Beta-x
SDO/HMI magnetograms	6696	8828	3646
SDO/HMI continuum images	5481	7993	2744

- (v) The selected FITS files must be of high quality. Corrupted data due to instrument failure or data of large background noise are omitted.

Based on the spatial distribution of magnetic polarities and penumbras, each AR magnetogram and continuum image is manually assigned to a corresponding magnetic type, respectively. In this labelling process, we have referred to the AR information provided by NOAA SRS files. It should be noted that a tiny amount of data have no obvious magnetic type information, which makes it difficult to label these samples manually. These data cannot be included in the analysis. The total amount of usable data is shown in Table 2. In all magnetograms, there are 6696 pictures of Alpha magnetic type, 8828 pictures of Beta, and 3646 pictures of Beta-x. In all continuum images, there are 5481 pictures of Alpha magnetic type, 7993 pictures of Beta, and 2744 pictures of Beta-x. The amount of data for each type in the data set is unevenly distributed. The smallest number of magnetic type is for Beta-x in continuum images.

The SDO/HMI magnetograms and continuum images data are read in batches, and the output is in PNG format according to

$$p_i = \frac{(255 + 0.9999) * (x_i - \min)}{\max - \min} \quad (1)$$

where  $\max$  and  $\min$  represent the maximum and minimum of the magnetic field strength, and  $x_i$  and  $p_i$  represent the original pixel value and processed pixel value of the same index in the data matrix, respectively. In order to maintain the uniformity of the data set, the gray value range of the magnetograms is processed. The value of  $\max$  is set to be 800.0, and the  $\min$  is set to be -800.0, that is, in the PNG images, the gray value of the pixel point with a magnetic field strength greater than or equal to 800 is assigned to 255, and the gray value of the pixel point with a magnetic field strength lower than or equal to -800 is assigned a value of zero. Figure 1 shows an example of AR magnetogram before and after processing. Compared to the unprocessed magnetogram, the processed magnetogram has more obvious characteristic information.

Besides, pictures obtained through the above process are interpolated into a uniform size of 160×80 to satisfy the input requirements by CNN structure.

## 4. Classification Model

*4.1. Training Model.* As a branch of artificial intelligence, machine learning algorithms allow machines to learn patterns from a large amount of historical data, thereby intelligently identifying new samples or predicting the future. The machine learning method adopted in this work is convolutional neural network (CNN), which is one of the typical algorithms of deep learning [30, 31]. CNN [32–35] is a kind of feedforward neural network with convolutional computation and deep structure, but, unlike the traditional multilayer feedforward neural network, the input of CNN is a two-dimensional pattern, images for example. The connection weight of CNN is a two-dimensional weight matrix, which is also called convolution kernel. Two basic operations in CNN are two-dimensional discrete convolution and pooling. Because CNN can process two-dimensional patterns directly, it has been widely used in the field of computer vision. A schematic of the CNN structure in this work is shown in Figure 2:

- (i) The input data sequentially pass through three layers of convolution. The size of the convolution kernel is 5×5. The number of first-layer convolution kernels is 32, and then each layer is doubled. The convolution kernel is used to extract image features. When performing convolution, the all-zero padding is used, and the step size is set to be 1.
- (ii) The result of each convolution is activated by the nonlinear activation function, followed by pooling. The number of features is controlled by the pooling to avoid overfitting. When performing pooling, the all-zero padding is used, the step size is set to be 2, and kernel size is 2×2.
- (iii) The convolved data features are fed into a fully connected neural network. Within the fully connected neural network, there is one hidden layer containing 512 neurons, and the output layer has 3 neurons, corresponding to three different magnetic types. Among the 3 nodes, the magnetic type corresponding to the node with the highest value is the recognition result of the model.
- (iv) The output of the fully connected neural network passes through a softmax function to obtain the probability distribution for the output classification. The result is then compared with the data label to get

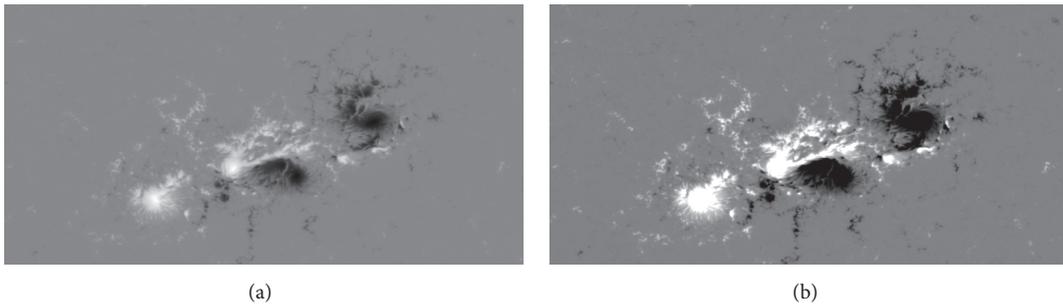


FIGURE 1: Magnetogram of NOAA 11158 AR (SDO/HMI, February 14, 2011, 22:24 UT). In gray images, white (positive polarity) and black (negative polarity) areas indicate strong magnetic fields, while gray areas show regions of weak magnetic fields. (a) and (b) are unprocessed and processed pictures, respectively.

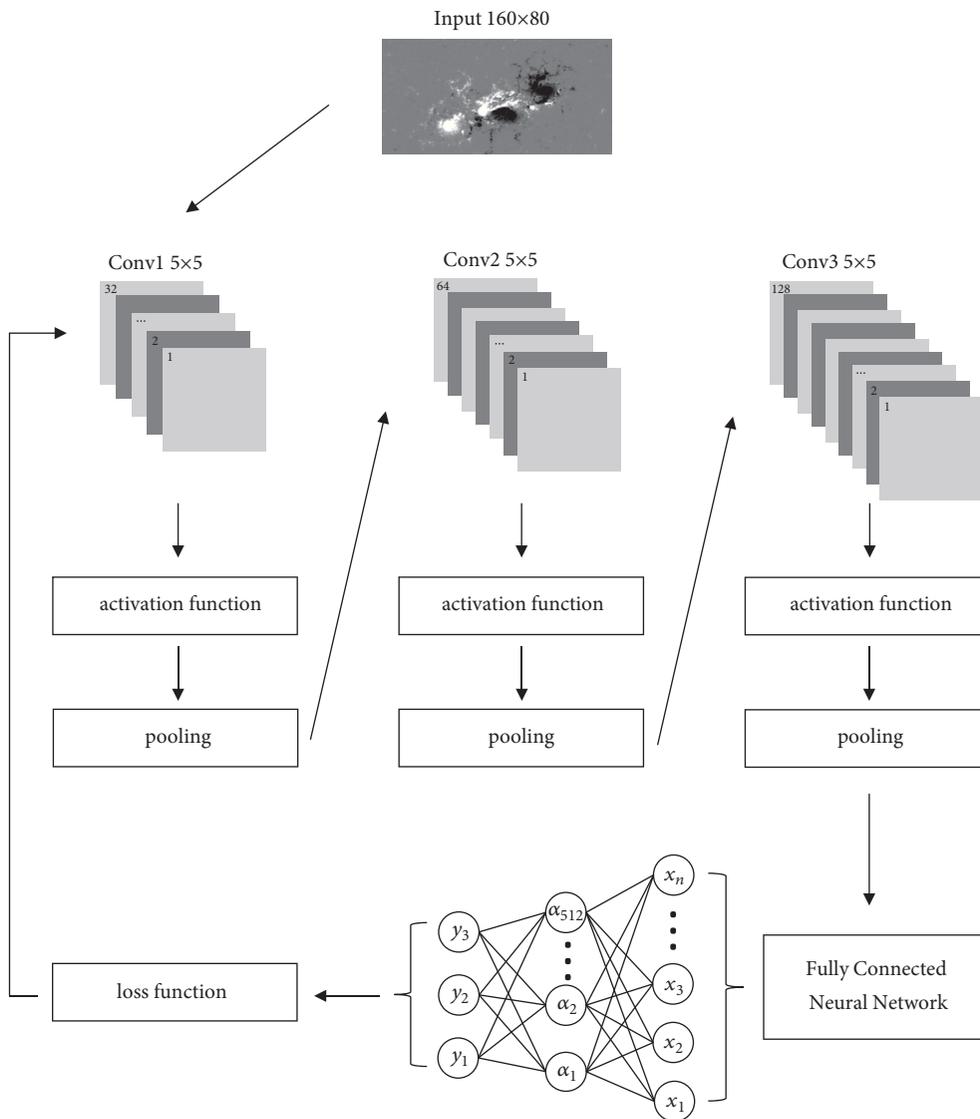


FIGURE 2: A schematic of the CNN structure.

TABLE 3: Performance of each training model.

Model	Data source	Loss function	Accuracy			
			Total	Alpha	Beta	Beta-x
A	magnetogram	5.53940	0.876667	0.958	0.763	0.903
B	continuum image	5.36543	0.954167	0.985	0.885	0.958
C	two-channel picture	5.45285	0.889167	0.958	0.793	0.913

TABLE 4: Three-class confusion matrix.

	Recognized Alpha	Recognized Beta	Recognized Beta-x
Observed Alpha	True Alpha(TA)	False Beta_A(AB)	False Beta-x_A(AC)
Observed Beta	False Alpha_B(BA)	True Beta(TB)	False Beta-x_B(BC)
Observed Beta-x	False Alpha_C(CA)	False Beta_C(CB)	True Beta-x(TC)

the cross-entropy, so as to obtain the loss function. The softmax function and the cross-entropy formula are given by the following:

$$\text{Softmax}(y_k) = \frac{e^{y_k}}{\sum_{j=1}^n e^{y_j}} \quad (2)$$

where  $n$  is the number of types, and  $y_k$  is the output of the  $k$ th network output unit.

$$H(Y_-, Y) = - \sum Y_- * \log Y \quad (3)$$

$Y_-$  is the truth value of the data label, and  $Y$  is the output of neural networks.

- (v) Optimize the model by minimizing the loss function. Each model stops training after 50,000 iterations.

**4.2. Results.** We have constructed three different classification models, namely, A, B, and C, respectively. Model A uses HMI magnetogram data as input, while model B uses HMI continuum image data. Model C uses both magnetogram and continuum images as two-channel input in CNN. The corresponding convolution kernel is also set to two channels in model C.

When training our models, it is required that the amount of data used for training or testing is the same for each class. At the same time, the combination of the magnetogram and continuum image in model C requires integration and matching of the two data sets. In order to satisfy these requirements, we have used the undersampling method to randomly select 2600 pictures for each class as inputs: 400 for testing and 2200 for training. Totally, there are 6600 pictures used for training and 1200 pictures for testing. The test set data do not participate in the training process.

The performance of all models is evaluated by the loss function and the accuracy rate. The loss function maps the value of a random event or its associated random variable to a nonnegative real number to represent the “risk” or “loss” of the random event. In applications, the loss function is often associated with the optimization problem as a learning criterion, which is to solve and evaluate the model by minimizing the loss function [36]. The total accuracy rate is

obtained by randomly extracting some samples from the test set for magnetic type recognition after the model training is completed. These sample data are not used in the training process and can better reflect the recognition performance of the model. The recognition accuracy for each magnetic type is obtained by traversing their test set. The results are shown in Table 3. In the three models, the value of loss function for model B is the smallest, and that of model A is the largest. The total accuracy of model B is the highest, exceeding 95%, and model A has the lowest total accuracy of about 87%. Comparing the performance of the three models in each class, the accuracy of model B is the highest. Although model A and C have the same recognition accuracy for Alpha, the accuracy of model C for Beta and Beta-x are better than that of model A. Among the three categories, Alpha has the highest accuracy while Beta has the lowest accuracy. This phenomenon emerges in all three training models.

In addition, the confusion matrix is used to evaluate the classification performance of the model for the magnetic type of the AR. The confusion matrix definition is shown in Table 4. The number of samples that are correctly recognized as “Alpha” is true Alpha (TA), the number of Alpha samples that are wrongly recognized as “Beta” is false Beta\_A (AB), and the number of Alpha samples that are wrongly recognized as “Beta-x” is false Beta-x\_A (AC). The number of samples that are correctly recognized as “Beta” is true Beta (TB), the number of Beta samples that are wrongly recognized as “Alpha” is false Alpha\_B (BA), and the number of Beta samples that are wrongly recognized as “Beta-x” is false Beta-x\_B (BC). The number of samples that are correctly recognized as “Beta-x” is true Beta-x (TC), the number of Beta-x samples that are wrongly recognized as “Alpha” is false Alpha\_C (CA), and the number of Beta-x samples that are wrongly recognized as “Beta” is false Beta\_C (CB). The classification result of the above three models is detailed in Tables 5–7. It can be seen from the tables that Alpha magnetic type is misidentified as Beta magnetic type occasionally when the model recognition error occurs. Beta magnetic type is sometimes misclassified as either Alpha or Beta-x, and the error rate is more than 4%. Beta-x magnetic type is mainly misclassified as Beta. When classifying manually, the disagreement due to individual experience and personal

TABLE 5: Details of Model A Classification Results.

Manual classification	Automatic recognition		
	<i>Alpha</i>	<i>Beta</i>	<i>Beta-x</i>
<i>Alpha</i> (num/400)	383 (95.8%)	17 (4.3%)	0 (0.0%)
<i>Beta</i> (num/400)	36 (9.0%)	305 (76.3%)	59 (14.8%)
<i>Beta-x</i> (num/400)	10 (2.5%)	29 (7.3%)	361 (90.3%)
<i>Total</i>	429	351	420

TABLE 6: Details of Model B Classification Results.

Manual classification	Automatic recognition		
	<i>Alpha</i>	<i>Beta</i>	<i>Beta-x</i>
<i>Alpha</i> (num/400)	394 (98.5%)	6 (1.5%)	0 (0.0%)
<i>Beta</i> (num/400)	18 (4.5%)	354 (88.5%)	28 (7.0%)
<i>Beta-x</i> (num/400)	4 (1.0%)	13 (3.3%)	383 (95.8%)
<i>Total</i>	416	373	411

TABLE 7: Details of Model C Classification Results.

Manual classification	Automatic recognition		
	<i>Alpha</i>	<i>Beta</i>	<i>Beta-x</i>
<i>Alpha</i> (num/400)	383 (95.8%)	17 (4.3%)	0 (0.0%)
<i>Beta</i> (num/400)	37 (9.3%)	317 (79.3%)	46 (11.5%)
<i>Beta-x</i> (num/400)	11 (2.8%)	24 (6.0%)	365 (91.3%)
<i>Total</i>	431	358	411

opinion is inevitable, mainly in distinguishing Beta magnetic type and Beta-x magnetic type.

True positive rate (TP rate) and false positive rate (FP rate) are defined to measure the performance of models. The TP rate is the percentage of positive samples correctly classified, and the FP rate is the percentage of negative samples that are misclassified. The formulas for calculating TP rate and FP rate of Alpha are shown in the following:

$$TP\ rate = \frac{TA}{TA + AB + AC} \quad (4)$$

$$FP\ rate = \frac{BA + CA}{BA + TB + BC + CA + CB + TC} \quad (5)$$

Using TP rate as the longitudinal axis and FP rate as the horizontal axis, Receiver Operating Characteristic (ROC) curves are drawn [37]. As shown in Figure 3, the (0, 1) point means that all samples are classified correctly. The nearer to (0, 1), the higher the classification accuracy of samples. Models A, B, and C are represented by red, green, and blue, respectively. The area of the quadrilateral formed by each point and (0, 0), (1, 1), (1, 0) points is the area under the curve (AUC) of that point. The larger the AUC value, the better the classification effect of the model. It can be seen more intuitively from Figure 3 that model B has the best classification performance among the three models, and Alpha has the best classification effect among the three kinds of magnetic types.

## 5. Conclusion and Discussion

In order to develop the automatic identification for the AR magnetic type based on the Mount Wilson classification scheme, we adopt CNN to train the SDO/HMI magnetogram and continuum image data from 2010 to 2017. We have constructed three models: A, B, and C, using magnetograms, continuum images, and the two-channel pictures as input, respectively. The conclusions are as follows. First, CNN has a productive performance in the identification of the AR magnetic types. The overall accuracy of all three models is over 87%, and the highest total accuracy is more than 95%. The recognition accuracy for Alpha type reaches 98%, and the accuracy for Beta type maintains above 88%. Second, the best recognition performance appears when continuum image is the sole input, followed by two-channel picture as input, and the performance of training model is relatively the worst when feeding only the magnetogram data. Finally, Alpha type is the easiest to recognize while the accuracy for Beta type is the lowest in the three categories. This phenomenon exists in all three training models.

By analyzing Mount Wilson sunspot magnetic types in a large amount of continuum images, it is found that the Alpha class is generally presented in a single sunspot, Beta class generally distributes in a relatively discrete and regular sunspot group, and the majority of Beta-x class appears in a relatively large and irregularly distributed sunspot group. Although the magnetic polarity, umbra, and penumbra information of sunspot groups need to be used at the same time in Mount Wilson sunspot classification scheme theoretically, it

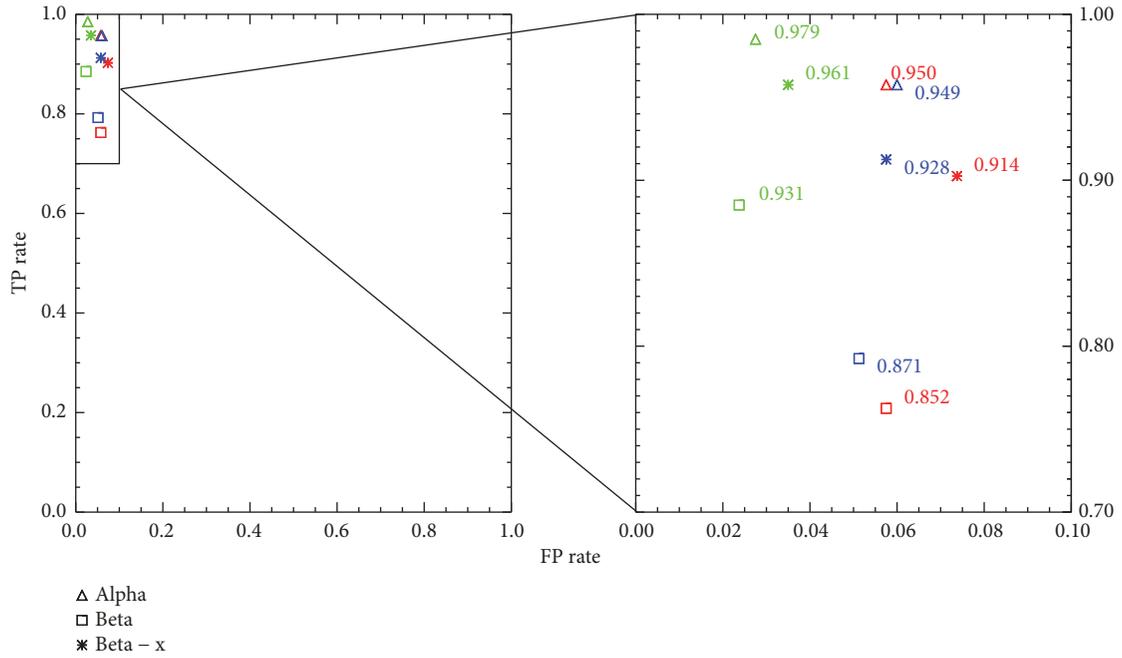


FIGURE 3: TP rate and FP rate of Alpha, Beta, and Beta-x in model A (red), model B (green), and model C (blue). The AUC value is marked beside each point.

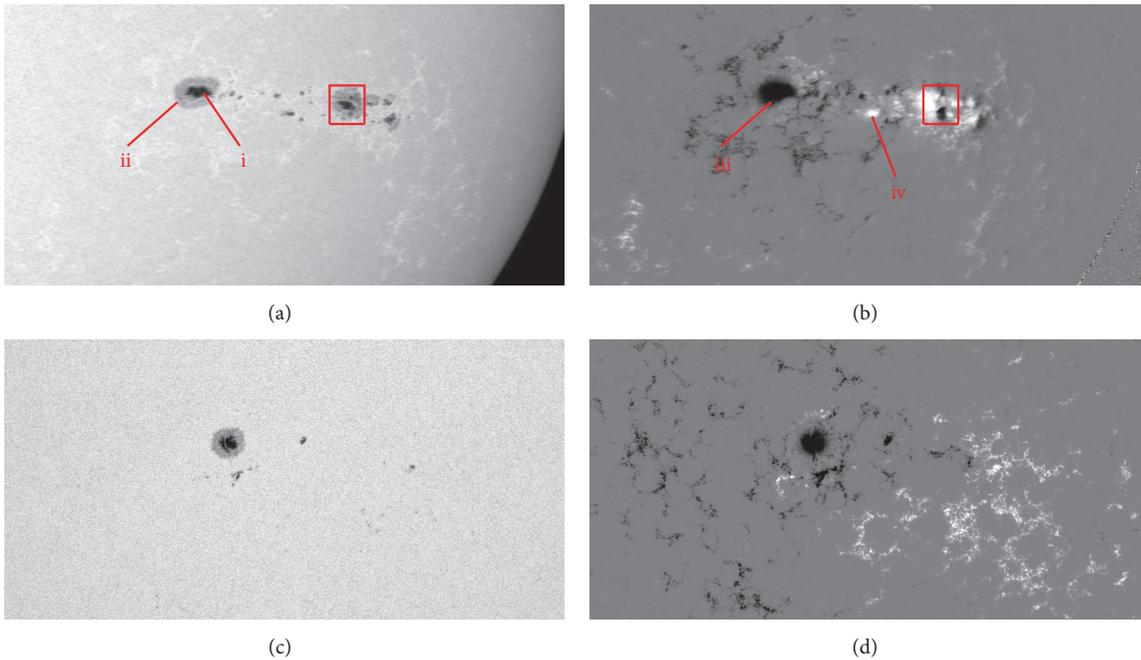


FIGURE 4: Two cases of misidentification of Beta-x as Beta in model B. (a) and (c) are the continuum images, whereas (b) and (d) are the magnetograms of the same ARs. In continuum images, the dark center is the umbra (i) and a lighter black part around umbra is the penumbra (ii). In magnetograms, the dark areas indicate the negative polarity (iii) region, and the bright areas indicate the positive polarity (iv) region.

is also feasible to use the continuum images alone to classify the three types that are defined in this work. However, it is confusing that the training performance with magnetogram or two-channel picture as input is not as good as that using continuum image only. Specific case analysis may shed a stripe of light over the black box. As shown in Figure 4,

for model B, the cases that the Beta-x magnetic type is misclassified into Beta magnetic type mainly consist of two different origins: (1) a heteropolar region can be seen in the magnetogram while there is a penumbra in the continuum image corresponding to the opposite polarity region, but no obvious multiple umbrae, as shown by the red box in Figures

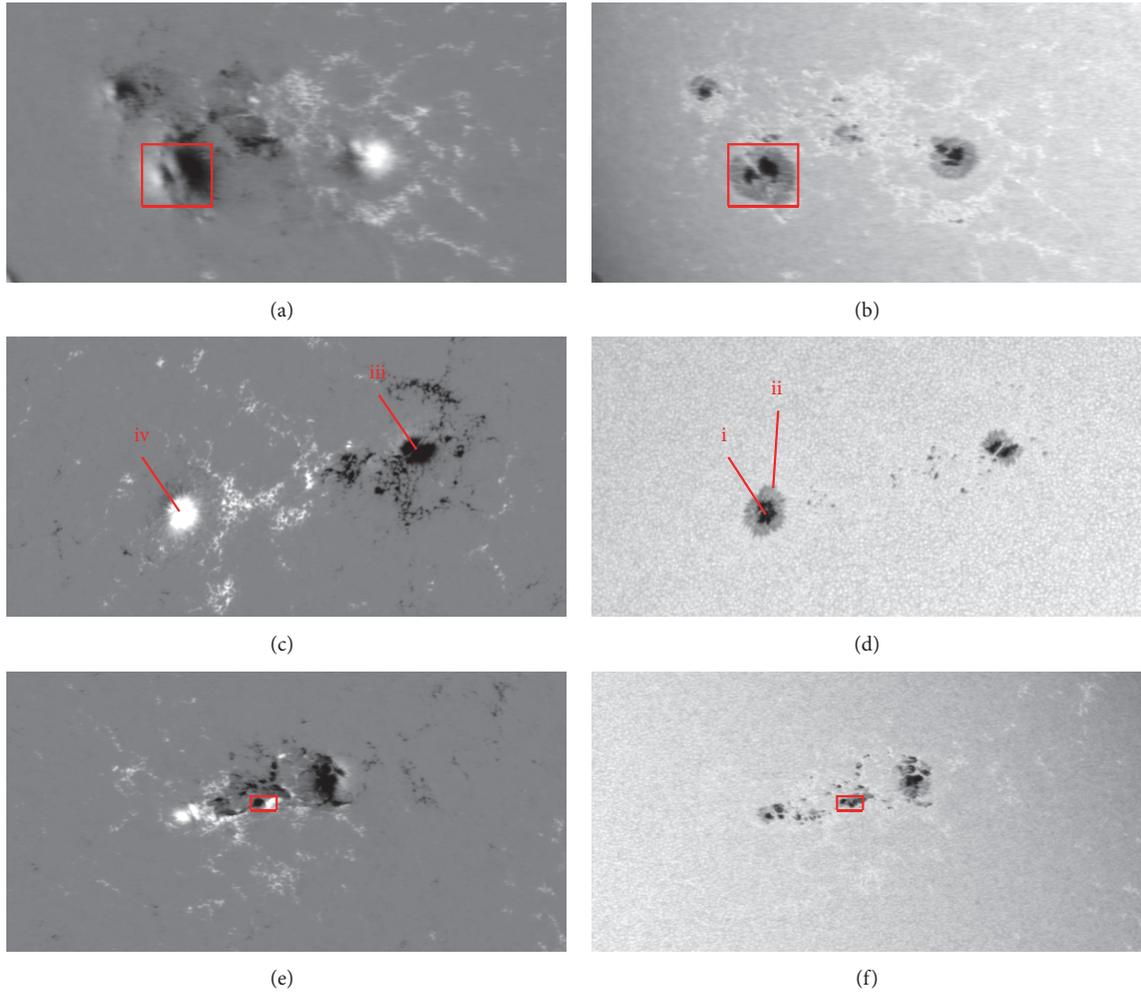


FIGURE 5: Three cases of misidentification of Beta-x as Beta in model A. (a), (c), and (e) are the magnetograms, whereas (b), (d), and (f) are the continuum images of the same ARs. In continuum images, the dark center is the umbra (i) and a lighter black part around umbra is the penumbra (ii). In magnetograms, the dark areas indicate the negative polarity (iii) region, and the bright areas indicate the positive polarity (iv) region.

4(a) and 4(b), mainly due to the inability to obtain the polarity information of the magnetic field in the continuum image; (2) there is no obvious heteropolar region in the magnetogram while the corresponding region in the continuum image has no obvious umbra, and the distribution of sunspots is diffusing, as shown in Figures 4(c) and 4(d). The latter may belong to the early or late stage of an AR. Such a situation is highly dependent on the subjectivity of manual prelabeling. In fact, although the classification rules are clarified, it is very difficult for the classification results to be 100% unified when experts perform manual classification.

As shown in Figure 5, for model A, there are three perspectives that the Beta-x magnetic type is misclassified into Beta magnetic type. First, as shown by the red box in Figures 5(a) and 5(b), there is no noticeable heteropolar region in the magnetogram, but multiple umbrae in the same penumbra can be seen in the corresponding continuum image. Errors may occur due to the inability to accurately obtain the umbra and penumbra information of the sunspot from the

magnetogram alone. Second, as shown in Figures 5(c) and 5(d), similar to the reason (2) of model B, the magnetic type of the AR is controversial and cannot be completely counted as a model classification error. The artificial subjective consciousness accounts for a large part of the reason. Third, the AR with sufficiently complex distribution and no obvious boundary between opposite polarities contains a bipolar sunspot group with opposite polarity umbrae sharing one penumbra that spans less than 2 degrees with respect to the solar center, as shown by the red box in Figures 5(e) and 5(f). The magnetic type Beta-Gamma-Delta can be determined from the magnetogram and continuum image by referring to the above Mount Wilson sunspot classification scheme. In fact, it can be clearly seen from the magnetogram alone that the magnetic type of the AR is Beta-x, but the model classification is wrong, which may be caused by the incomplete feature extraction. This is also the main reason for the poor training performance when the model uses only the magnetogram data.

The structure of the magnetogram is more complex than that of continuum image, but the amount of data available for training is insufficient, yielding that the CNN method is not able to extract picture features well. This is possibly one of the main reasons why the training performance is not good when using only the magnetogram data or using two-channel picture as the input.

In further work, we plan to improve the recognition accuracy and try to refine the classification. Due to the different information content of magnetograms and continuum images, we will use different convolutional network structures to extract feature information for these two kinds of data separately. Then the extracted features can be integrated and fed into a fully connected neural network. Furthermore, it is necessary to continuously supplement and balance the data set.

## Data Availability

The magnetograms and continuum images data used to support the findings of this study are observed by SDO/HMI; all of the FITS files we have used are downloaded from the website <http://jsoc.stanford.edu/>. The SRS text files can be downloaded from the NOAA/SWPC website <https://www.swpc.noaa.gov/>. The train set and test set data for machine learning of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

The data used in this study are courtesy of the SDO/HMI and NOAA/SWPC science teams. This work is supported by the Beijing Municipal Science and Technology Project (Project no. Z181100002918004) and the National Defense Science and Technology Innovation Special Zone.

## References

- [1] G. E. Hale, "The zeeman effect in the sun," *Publications of the Astronomical Society of the Pacific*, vol. 20, no. 123, pp. 287-288, 1908.
- [2] K. Sakurai, "Motion of sunspot magnetic fields and its relation to solar flares," *Solar Physics*, vol. 47, no. 1, pp. 261-266, 1976.
- [3] P. O. Taylor, "The relationship between sunspot and solar flare activities for the period 1974.6 - 1987.6," *Journal of the American Association of Variable Star Observers*, vol. 17, 1987.
- [4] Z. Shi and J. Wang, "Delta-sunspots and X-class flares," *Solar Physics*, vol. 149, no. 1, pp. 105-118, 1994.
- [5] I. Sammis, F. Tang, and H. Zirin, "The dependence of large flare occurrence on the magnetic structure of sunspots," *The Astrophysical Journal*, vol. 540, no. 1, pp. 583-587, 2000.
- [6] Y. Cui, R. Li, L. Zhang, Y. He, and H. Wang, "Correlation between solar flare productivity and photospheric magnetic field properties," *Solar Physics*, vol. 237, no. 1, pp. 45-59, 2006.
- [7] Y. Zhang, J. Liu, and H. Zhang, "Relationship between rotating sunspots and flares," *Solar Physics*, vol. 247, no. 1, pp. 39-52, 2008.
- [8] M. Hahn, S. Gaard, P. Jibben, R. C. Canfield, and D. Nandy, "Spatial relationship between twist in active region magnetic fields and solar flares," *The Astrophysical Journal*, vol. 629, no. 2 I, pp. 1135-1140, 2005.
- [9] M. Zhao, J. Chen, and Y. Liu, "Statistical analysis of sunspot groups and flares for solar maximum and minimum," *Scientia Sinica*, vol. 44, no. 1, pp. 109-120, 2014.
- [10] Y. Cui, R. Li, H. Wang, and H. He, "Correlation between solar flare productivity and photospheric magnetic field properties II. magnetic gradient and magnetic shear," *Solar Physics*, vol. 242, no. 1-2, pp. 1-8, 2007.
- [11] X. L. Yan, L. H. Deng, Z. Q. Qu et al., "The phase relation between sunspot numbers and soft X-ray flares," *Astrophysics & Space Science*, vol. 333, no. 1, pp. 11-16, 2011.
- [12] A. Raboonik, H. Safari, N. Alipour, and M. S. Wheatland, "Prediction of solar flares using unique signatures of magnetic field images," *The Astrophysical Journal*, vol. 834, no. 1, article no. 11, 2017.
- [13] A. E. McCloskey, P. T. Gallagher, and D. S. Bloomfield, "Flare forecasting using the evolution of McIntosh sunspot classifications," *Journal of Space Weather and Space Climate*, vol. 8, article A34, 14 pages, 2018.
- [14] T. Ataç, "Statistical relationship between sunspots and major flares," *Astrophysics and Space Science*, vol. 129, no. 1, pp. 203-208, 1987.
- [15] P. S. McIntosh, "The classification of sunspot groups," *Solar Physics*, vol. 125, no. 2, pp. 251-267, 1990.
- [16] S. Eren, A. Kilcik, T. Atay et al., "Flare-production potential associated with different sunspot groups," *Monthly Notices of the Royal Astronomical Society*, vol. 000, no. 19, 2017.
- [17] V. Zharkova, S. Ipson, A. Benkhalil, and S. Zharkov, "Feature recognition in solar images," *Artificial Intelligence Review*, vol. 23, no. 3, pp. 209-266, 2005.
- [18] S. Zharkov, V. Zharkova, S. Ipson, and A. Benkhalil, "Technique for automated recognition of sunspots on full-disk solar images," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 15, pp. 2573-2584, 2005.
- [19] S. Goel and S. K. Mathew, "Automated detection, characterization, and tracking of sunspots from SoHO/MDI continuum images," *Solar Physics*, vol. 289, no. 4, pp. 1413-1431, 2014.
- [20] S. H. Nguyen, T. T. Nguyen, and H. S. Nguyen, "Rough set approach to sunspot classification problem," in *Proceedings of the International Conference on Rough Sets*, Springer-Verlag, Germany, 2005.
- [21] T. T. Nguyen, C. P. Willis, D. J. Paddon, S. H. Nguyen, and H. S. Nguyen, "Learning sunspot classification," *Fundamenta Informaticae*, vol. 72, no. 1-3, pp. 295-309, 2006.
- [22] T. Nguyen, C. Willis, D. Paddon, and H. Nguyen, "A hybrid system for learning sunspot recognition and classification," in *Proceedings of the 2006 International Conference on Hybrid Information Technology*, pp. 257-264, IEEE Computer Society, Cheju Island, November 2006.
- [23] T. Colak and R. Qahwaji, "Automated mcintosh-based classification of sunspot groups using MDI images," *Solar Physics*, vol. 248, no. 2, pp. 277-296, 2008.
- [24] M. A. Abd, S. F. Majed, and V. Zharkova, "Automated classification of sunspot groups with support vector machines," in *Proceedings of the Technological Developments in Networking, Education and Automation*, pp. 321-325, 2010.

- [25] S. Padinhatteeri, P. A. Higgins, D. Shaun Bloomfield, and P. T. Gallagher, "Automatic detection of magnetic  $\delta$  in sunspot groups," *Solar Physics*, vol. 291, no. 1, pp. 41–53, 2016.
- [26] G. Hale E, F. Ellerman, and S. Nicholson B, "The magnetic polarity of sun-spots," *The Astrophysical Journal*, vol. 49, no. 3, p. 153, 1919.
- [27] W. D. Pesnell, B. J. Thompson, and P. C. Chamberlin, "The solar dynamics observatory (SDO)," *Solar Physics*, vol. 275, no. 1, pp. 3–15, 2012.
- [28] P. H. Scherrer, J. Schou, R. I. Bush et al., "The helioseismic and magnetic imager (HMI) investigation for the solar dynamics observatory (SDO)," *Solar Physics*, vol. 275, no. 1-2, pp. 207–227, 2012.
- [29] M. G. Bobra, X. Sun, J. T. Hoeksema et al., "The helioseismic and magnetic imager (HMI) vector magnetic field pipeline: sharps – space-weather hmi active region patches," *Solar Physics*, vol. 289, no. 9, pp. 3549–3578, 2014.
- [30] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [31] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *The American Association for the Advancement of Science: Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [32] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998.
- [33] C. Nebauer, "Evaluation of convolutional neural networks for visual recognition," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 9, no. 4, pp. 685–696, 1998.
- [34] J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.
- [35] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, vol. 1, MIT Press, Cambridge, Mass, USA, 2016.
- [36] P. Murugan, "Feed forward and backward run in deep convolution neural network," in *Proceedings of the 20th International Conference on Computer Vision and Image Processing*, 2017.
- [37] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

## Research Article

# Effect of Ap-Index of Geomagnetic Activity on S&P 500 Stock Market Return

Lifang Peng, Ning Li , and Jingwen Pan

*School of Management, Xiamen University, Xiamen, China*

Correspondence should be addressed to Ning Li; [lining08062012@163.com](mailto:lining08062012@163.com)

Received 24 February 2019; Revised 21 May 2019; Accepted 24 June 2019; Published 14 July 2019

Guest Editor: Liyun Zhang

Copyright © 2019 Lifang Peng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Geomagnetic activity with global influence is an essential object of space weather research and is a significant link in the section of the solar wind-magnetospheric coupling process. Research so far provides strong evidence that geomagnetic activity affects stock investment decisions by influencing human health, mood, and human behaviours. Therefore, this research investigates the empirical association between geomagnetic activity and stock market return. Overall, we find that geomagnetic activity exerts a negative influence on the return of the US stock market. Further, market liquidity effectively magnifies the effect of geomagnetic activity. Inconsistent with previous literature, this effect is not mainly caused by the semiannual variation of geomagnetic activity. Our research contributes to the introduction of geomagnetic indices to financial economics studies on the impact of geomagnetic activity influence on stock market return.

## 1. Introduction

The geomagnetic field is formed by the superposition of different magnetic fields from various sources. When the corresponding disturbance of the geomagnetic field occurs, the geomagnetic activities form. These disturbances vary in length and intensity and significantly affect human activities in terms of economics. Since the 1960s, people have been aware of the critical influence of cosmic activity, especially geomagnetic activities, on human behaviour. Friedman et al. (1963) initially explored the relationship between human health and geomagnetic activity parameters [1]. We summarize the research progress on the correlation between geomagnetic activity and the stock market into three ways: geomagnetic activity and human behaviours; human behaviours and the stock market; geomagnetic activity and the stock market.

First, geophysics research has provided evidence of a relationship between geomagnetic activity and human behaviours. These human behaviours can be concluded into three aspects: the influence of geomagnetic activity on the body's mood and nervous system; geomagnetic activity on human diseases; and geomagnetic activity on birth or death rate. The occurrence of geomagnetic activity affects the body's

emotional and nervous system [2–4]. Zakharov (2001) found that the effect is most marked during the recovery phase of geomagnetic storms and accompanied by the inhibition in the central nervous system [5]. According to Tarquini (1998), through influencing the activity of the pineal gland, geomagnetic activities cause imbalances and disruptions of the circadian rhythm of melatonin production [6], a factor that plays an important role in mood disturbances. Stoilova and Dimitrova (2008) examined blood pressure, heart rate, and electrocardiograms during perturbed and quiet days according to the Ap index; in particular, there was a clear tendency for changes in blood pressure during increased geomagnetic activity [7]. Mendoza (2010) studied that at middle and low latitudes there are biological consequences to the solar/geomagnetic activity coinciding [4], finding that geomagnetic perturbations cause gender differences, age differences, and myocardial infarctions (death or occurrence) influences.

Second, Economics and Psychology research also builds the relationship between human behaviours and the stock market, especially the correlation between mood and abnormal returns. Lerner et al. (2015) showed that mood affects economic decision making through a variety of pathways [8]. Shu (2010) showed that both equity and bill prices

correlate positively with investor mood [9], with higher asset prices associated with better mood, and conversely, expected asset returns correlate negatively with investor mood. These findings suggested that investor mood is a vital factor in equilibrium asset prices and returns. Moreover, the relationship between human behaviours and the stock market also exists in other aspects, like physical health [10].

Based on the above literature, it is ready to connect the geomagnetic activity with the stock market indirectly. Instead, some research also builds the direct relationship between geomagnetic activity and the stock market. Krivelyova & Robotti (2003) demonstrated lower stock market returns during periods of high geomagnetic activity and provided an explanation with investment mood as a mediated variable [11]. Belkin (2013) also found the strong direct connection between long stagflation waves in the USA and super solar cycles, and the strong inverse connection between seasonal geomagnetic storms and economic cycles in the US and Russian economics [10]. Besides, some research focused on the effect of weather on the stock market, using geomagnetic storms as one of the variables [9, 12–14]. However, these papers focus on the effect of extreme or irregular geomagnetic activities. To explore the day-to-day counterpart, instead of dealing with extreme and rare events such as geomagnetic storms, this research explores causality between geomagnetic activity and the stock market by using monthly US stock market indices and monthly levels of geomagnetic indices, Ap. SAD (Seasonal Affective Disorder) for the potential semiannual variation will be used in order to test whether the semiannual variation is one underlying cause for the causality.

This paper also explores how the effect of geomagnetic varies considering variations under two conditions. On one hand, when considering the variations of geomagnetic activities, semiannual variation was one of the earliest recognized patterns in geomagnetic activity [15, 16]. The semiannual variation in geomagnetic activity appears as spring and fall maximums in long-term averages of the various indices of geomagnetic activity. Kamstra et al. (2003) suggested that, in the process of the earth's seasonal alternation, the daily light time in a particular region will be different, and the length of the day and night changes will affect the internal rhythm of the human body [17], thus affecting the trading behaviour of investors. Therefore, we explore whether the targeted causality also has a semiannual effect similar to the geomagnetic activity itself.

On the other hand, motivated by Lanfear et al. (2018), abnormal illiquidity is only able to account for a small fraction of the observed abnormal returns caused by extreme weather events [18]. We then explore whether, during day-to-day counterparts, liquidity can still have a significant influence on the correlation between geomagnetic activities and the stock market indices. We find that market liquidity is positively related to the geomagnetic effect on the stock market. This finding is consistent with Krivelyova & Robotti (2003); the geomagnetic storm's effect is more significantly stronger in markets in worldwide important and open economies [11].

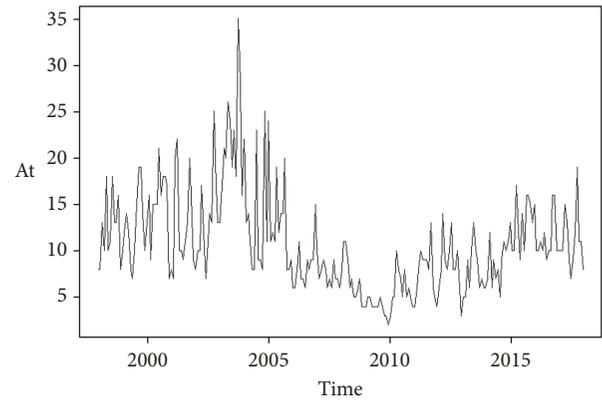


FIGURE 1: Overview of monthly DGD (1998-2017).

Our research implicates in two ways. First, by focusing on the liquidity and periodic output, this paper provides support for the two conditions. Second, we introduce a geomagnetic relationship to the financial economics literature. The rest of the paper is organized as follows. The data description and the methodology used for research questions are reported in Section 2. In Section 3, we present the results related to causality between geomagnetic activity and the stock market. In Section 4, we present the findings related to semiannual and liquidity variation. We conclude the whole paper in Section 5.

## 2. Materials and Methods

We use an empirical study for this research. The vast majority of empirical studies on daily geomagnetic data (DGD) use either the Ap or the Kp index to capture the intensity of the environmental magnetic field. In this paper, we choose the monthly Ap index as a proxy for DGD, named as  $DGD_t$ . The chosen period is from January 1998 to December 2017. Geomagnetic data for the Earth were obtained from the Space Weather Prediction Center, which is a part of the National Oceanic & Atmospheric Administration (NOAA). Figure 1 shows an overview of monthly DGD. Figure 2 shows an annual variation of the Ap index. As discussed above, there are significantly spring and fall maximums in long-term averages of this index.

We use S&P 500 as the index to describe the US stock market, named as  $SMI_t$ . The S&P 500 Index (formerly Standard & Poor's 500 Index) is a market-capitalization-weighted index of the 500 largest US publicly traded companies by market value. The index is widely regarded as the best single gauge of large-cap US equities. Monthly S&P 500 are obtained from Yahoo Finance. Figure 3 shows an overview of the monthly S&P 500. Figure 4 shows annually variation of S&P 500. Intuitively, the trend and seasonal changes of  $SMI_t$  and  $DGD_t$  are different.

**2.1. Augmented Dickey-Fuller Test.** Testing time series data for stationarity is a prerequisite for moving forward since the presence of unit roots would lead to the regression

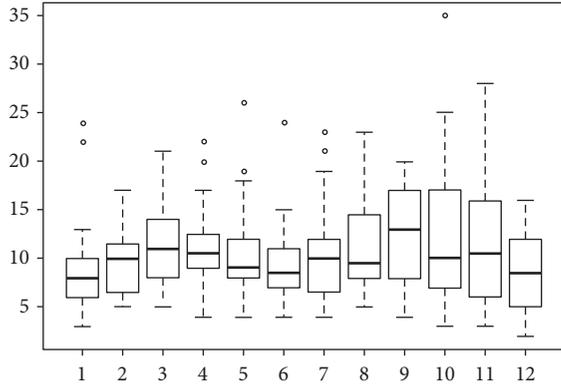


FIGURE 2: Annual variation of monthly DGD (1998-2017).

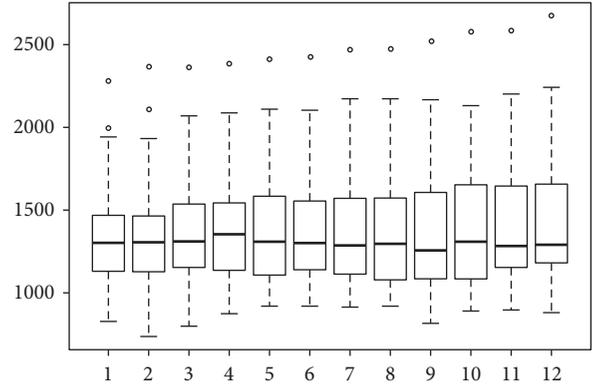


FIGURE 4: Annual variation of monthly S&amp;P 500 (1998-2017).

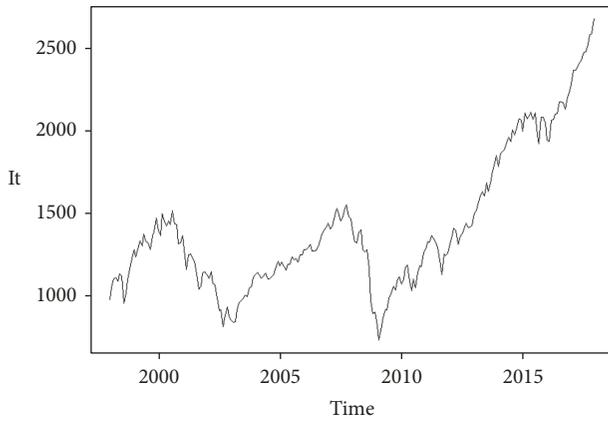


FIGURE 3: Overview of the monthly S&amp;P 500 (1998-2017).

being spurious unless there is the existence of at least one cointegrating relationship. In order to check the stationarity of the variables that are considered in this study, we use the Augmented Dickey-Fuller (ADF) stationarity test to detect the possible existence of unit roots in the data set. The variables should ideally be stationary at either  $I(0)$  or their first difference forms,  $I(1)$ . Once the variables are found to be stationary, the cointegration test is to be followed.

**2.2. Engle-Granger Test.** The most well-known test for cointegration, suggested by Engle and Granger (1987), is to run a static regression [19] (after first having verified that  $y_t$  and  $x_t$  both are  $I(1)$ ) of  $y_t = \theta x_t + e_t$  and  $\hat{e}_t = y_t - \hat{\theta}x_t$ , where  $x_t$  is one- or higher-dimensional. The asymptotic distribution of  $\theta$  is not standard, but the test suggested by Engle and Granger was to estimate  $\hat{\theta}$  by OLS and the test for unit roots in  $\hat{e}_t$ .

**2.3. Granger Causality Test.** According to cointegration analysis, when two variables are cointegrated, then there exists at least one direction of causality. Granger causality [20], introduced by Granger (1969), is one of the important matters that have been much studied in empirical macroeconomics and empirical finance. Only when the variables are cointegrated, it is possible to deduce that a long-run relationship exists

between the nonstationary time series. When we take  $y$  and  $x$  as the variables of interest, then the Granger causality test (Granger (1969)) determines whether past values of  $y$  add to the explanation of current values of  $x$  as provided by the information in past values of  $x$  itself. If previous changes in  $y$  do not help explain current changes in  $x$ , then  $y$  does not Granger cause  $x$ . The equations are as follows:

$$y_t = \alpha + \sum_{i=1}^t \beta_i y_{t-i} + \varepsilon_{t,1} \quad (1)$$

$$y_t = \alpha + \sum_{i=1}^t \beta_i y_{t-i} + \sum_{j=1}^t \gamma_j x_{t-j} + \varepsilon_{t,2} \quad (2)$$

Similarly, we can examine if  $x$  Granger causes  $y$  to just be interchanging them and carrying out this process again. There could be four probable outcomes: (i)  $x$  Granger causes  $y$ ; (ii)  $y$  Granger causes  $x$ ; (iii) both  $x$  and  $y$  Granger cause the other; (iv) neither of the variables Granger causes the other. We consider the above set of equations for all possible pairs of ( $SMI_t$ ,  $DGD_t$ ) series in the group.

**2.4. Regression.** As shown in Table 1, there are some other variables which would be used in the following empirical tests. Specifically, ILLIQ is the measure of liquidity of the stock market. The bigger the ILLIQ is, the smaller the stock market liquidity is.

After Granger causality test, we use linear regression as a robust test for the relationship between geomagnetic activity and the stock market, as shown in (3). We predict that  $\beta_1$  is significantly negative:

$$SMI_t = \beta_0 + \beta_1 DGD_t + \varepsilon_t \quad (3)$$

$$SMI_t = \beta_0 + \beta_1 DGD_t + YEAR + \varepsilon_t \quad (4)$$

To test seasonal variations in the targeted relationship, we consider the  $SAD$  index to represent this variation. In (5),  $SAD$  captures the effect of  $DGD$  varying semiannually, and  $\varepsilon_t$

TABLE 1: Variable Specifications.

Measure	Description
DGD	Monthly Geomagnetic activity index (Ap). The data source is Space Weather Prediction Center, a part of NOAA. Data period is Jan. 1998 – Dec. 2017.
SMI	Monthly S&P 500, popular US stock market indices. The data source is Yahoo Finance. Data period is Jan. 1998 – Dec. 2017.
SAD	Seasonal Affective Disorder index. Dummy variable. Variable equals to 1 if the set of indexes occurs in Jan., Feb., Mar., Nov. or Dec., and zero otherwise.
$p_H$	The monthly highest price in S&P 500.
$p_L$	The monthly lowest price in S&P 500.
VOLD	The stock trade-off volume among the 500 largest US publicly traded companies per month.
ILLIQ	The market illiquidity among the 500 largest US publicly traded companies per month. Compute the stocks' illiquidity as the ratio of dollar volume to the price difference. $ILLIQ_t = (p_{H,t} - p_{L,t})/VOLD_t * 10^{10}$ .
YEAR	Period Dummy variable for years 1998 to 2017. For example, macroeconomic environment changes would happen year by year, and this index could control this kind of effect.

captures the remaining part. We predict that  $\beta_1$  is significantly negative:

$$DGD_t = \alpha_0 + \alpha_1 SAD_t + e_t \quad (5)$$

$$SMI_t = \beta_0 + \beta_1 SAD_t + \beta_2 e_t + YEAR + \varepsilon_t \quad (6)$$

During a hurricane, assets with higher liquidity suffer more loss than assets with lower liquidity, such as gold [18]. Motivated by Amihud (2002), we consider the illiquidity index to represent this variation [21]. Note that ILLIQ represents the market illiquidity level per month. Therefore, for the third question, here we predict an adverse effect of this variable on the relationship between the geomagnetic index and stock market indices which is presented by (7). Some economic shocks may differentially affect the stock market with different degrees of reliance on geomagnetic activities. Therefore, interactions between  $DGD$  and  $ILLIQ$  on  $SMI$  variables are included to help control for any preexisting noises.  $\beta_3$  is the coefficient of the interaction. We predict that this coefficient is significantly negative. To keep the comparability of  $\beta_1$  in (3) and (7), motivated by Balli, Sørensen (2013), we further modify the model as (8), where  $\overline{DGD}_t$ ,  $\overline{ILLIQ}_t$  are the average value of the corresponding variables [22]:

$$SMI_t = \beta_0 + \beta_1 DGD_t + \beta_2 ILLIQ_t + \beta_3 DGD_t * ILLIQ_t + YEAR + \varepsilon_t \quad (7)$$

$$SMI_t = \beta_0 + \beta_1 DGD_t + \beta_2 ILLIQ_t + \beta_3 (DGD_t - \overline{DGD}_t) * (ILLIQ_t - \overline{ILLIQ}_t) + YEAR + \varepsilon_t \quad (8)$$

Table 2 shows the statistical summary of these variables used in the above equations. Table 3 shows the correlation matrix. The result shows a significantly negative relation between geomagnetic activity index and stock market indices as predicted. Also, we find a robust negative correlation which means that the market illiquidity varies with time. One explanation is that this variable is measured as monthly data.

However, we further test the correlation between  $ILLIQ$  and month dummies and find no strong correlation (correlation = 0.01, p-value = 0.87). Then we rule out this possibility. Another explanation is the variance of the macroeconomic environment. For example, the promulgation of regulations and laws would have implications for the present and future years.

### 3. Causality between Geomagnetic Activity and the Stock Market

**3.1. Main Test.** The first question is whether geomagnetic activity is negatively related to the US stock market. For this problem, we use the Granger Causality Test to testify the causality between geomagnetic activity and the stock market and use linear regression as robustness for the magnitude of the effect.

Results from the ADF test are provided in Table 4. The results for the ADF unit root tests for variables at the primary level (non-first-order difference). As shown in the first two rows in Table 4, results are not stable. Then we further test whether SMI and DGD are cointegrated. The cointegration result of the primary level, as shown in the third row in Table 4, rejects the null hypothesis at 1% significant level. So, we can show that this regression is not spurious. The results from the ADF test confirm that both of the variables are stationary at their first differenced forms, and our following tests use the first-differenced data.

As two time-series variables are stationary and cointegrated, we continue the Granger Causality Test augmented with a lagged error correction term if the series are cointegrated. We use AIC and SC information law to choose the optimal lag length which equals 1. The p-value on the lagged explanatory variables of error correction indicates the significance of the short-run causal effects. Beginning with the short-run effects, we find that there are unidirectional causalities from DGD to SMI; in other words, SMI does not Granger cause DGD, but DGD Granger causes SMI. Additionally, we do the test in the case where lag order equals 2 and find a similar conclusion. This unidirectional

TABLE 2: Summary of Statistics.

	Mean	Std.Dev	Maximum	Minimum	Observations
DGD	10.76	5.23	35.00	2.00	240
SMI	1414.73	417.95	2673.61	735.09	240
P <sub>H</sub>	1452.19	413.33	2694.97	832.98	240
P <sub>L</sub>	1360.93	414.20	2605.52	666.79	240
VOLD(*10 <sup>-10</sup> )	5.90	3.21	16.20	1.15	240
ILLIQ	22.63	20.49	109.15	4.2	240

TABLE 3: Correlation Matrix.

	DGD	SMI	SAD	ILLIQ	YEAR
DGD	1.00				
SMI	-0.13**	1.00			
SAD	-0.04	0.06	1.00		
ILLIQ	0.08*	-0.11	-0.01	1.00	
YEAR	-0.01	0.04	0.00	-0.71***	1.00

Note: \*\*\*, \*\*, and \* denote statistically significant at 1%, 5%, and 10% levels, respectively. See Table 1 for variable definitions.

TABLE 4: Augmented Dickey-Fuller (ADF) Test Results.

	Variable	Dickey-Fuller	Lag	p-value	Conclusion
Unit root test for primary level	SMI	-0.66	6	0.9731	Not Stationary
	DGD	-2.75	6	0.2618	Not Stationary
Cointegration test		-5.55	6	<0.01	Stationary
Unit root test for 1 <sup>st</sup> differenced	SMI	-7.84	6	<0.01	Stationary
	DGD	-5.57	6	<0.01	Stationary

causality result provides statistical support for our primary prediction that geomagnetic activity is negatively related to the US stock market. Furthermore, geomagnetic activity has a unidirectional effect on the stock market.

**3.2. Robustness.** To testify the magnitude and direction of the targeted relationship, we also test the relation between geomagnetic activity and the stock market by linear regression. Table 6 reports the regression results for (3). Column (3) is without control variable year dummy, while column (4) is for regression with controls. Consistent with the causality results reported in Table 5, results in Table 6 are both significantly negative, suggesting that there is indeed an inverse relationship between geomagnetic activity and U.S. stock market (p-value < 0.05). Further, it is indicated that the effect of geomagnetic activity on the stock market is economically significant. The predicted possibility of an inverted market signal is 1.2%. Comparing the results in column (3) and column (4), the control variable YEAR does not have an evident influence on the conclusion.

## 4. Additional Tests

**4.1. Variation in Semiannual Periods.** Based on the above results, we hold a further question of whether the targeted causality also has a semiannual variation similar to the

geomagnetic activity itself. We use Two-Stage Least Squares (2SLS) Regression. In the first stage, we regress SAD on DGD and decompose DGD into two parts. One represents the variations in winter and summer, and the other is the remaining part (4). In the second stage, we regress the two parts on SMI. As shown in Table 7, we do not find significant results to support our second prediction (5). We find that the significant effect of DGD on SMI is mainly due to the remaining part without seasonal variations, other than the semiannual variation effect as predicted. To some extent, this is also consistent with Dowling, M., & Lucey (2008) [13]. Although they prefer to use seasonal affective disorder to explain changes in investors' mood, most of them do not find significant results. We also use month dummy variable (data in January equals 1, in February equals 2, ..., and in December equals 12) or half year dummy variable (data in January–June equals 1 and in July–December equals 0) for the similar test; however, we do not find expected results either.

One explanation is that monthly aggregate data remove some difference among daily or smaller-unit data and cause too much noise. Another potential explanation is that the trend of increasing geomagnetic activity which removed seasonality effect is the primary cause of the targeted causality. Our results support the second explanation.

**4.2. Variation in Liquidity.** Furthermore, we explore whether the relationship between geomagnetic activity and the US

TABLE 5: Granger Causality Test Results.

Null Hypothesis	Lag	p-value	Conclusion
DGD does not Granger cause of SMI	1	0.0505	Causality significant at 10% level
SMI does not Granger cause of DGD	1	0.2012	No Causality
DGD does not Granger cause of SMI	2	0.0233	Causality significant at 5% level
SMI does not Granger cause of DGD	2	0.5333	No Causality

TABLE 6: Regression Results for Equ. (3) and (4).

	SMI			
	(3)		(4)	
	Coefficient	p-value	Coefficient	p-value
Intercept	0.0000	1.000	-0.6509	0.5026
DGD	-0.0012	0.0480**	-0.0012	0.0489**
YEAR	No		Yes	
Adj. R-squared	0.0122		0.0099	
Obs.	240		240	

TABLE 7: Regression Results for Equ. (5), (6) and (8).

	DGD		SMI		SMI	
	(5)		(6)		(8)	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Intercept	0.1476	0.7020	-0.0023	0.5216	0.0035	0.3939
SAD	-0.3543	0.5530	0.0056	0.3263		
e			-0.0012	0.0523*		
DGD					-0.0007	0.1923
ILLIQ					-0.0001	0.3811
DGD*ILLIQ					-0.0001	0.0009* * *
Adj. R-squared	-0.0027		0.0115		0.0585	
Obs.	240		240		240	

stock market would be moderated by market liquidity. As mentioned before, ILLIQ is the inverse index for market liquidity; in other words, the bigger the ILLIQ is, the smaller the stock market liquidity is. Therefore, the coefficient of DGD\*ILLIQ on SMI in Table 7 column (6) is negative which means that market liquidity is positively related to the geomagnetic effect on the stock market as predicted. This finding is also consistent with the findings in Lanfear et al. (2018) [18].

One potential explanation for the results about variation in market liquidity is that they are influenced by geomagnetic waves subconsciously or unconsciously; investors suffer bad moods which lead to a preference for more meaningful decisions and an increase in risk aversion [23, 24]. Meanwhile, market liquidity would magnify this kind of risk estimation which causes a higher premium on equity with higher liquidity.

What is more, it is evident that after adding the market illiquidity, R-squared has increased from 1.22% in Table 6 column (3) to 5.85% in Table 7 column (6) and the significance of regression results has been strengthened simultaneously. This change provides evidence for our prediction that monthly aggregate data covered up some variations in the daily level.

However, the fact that with so much noise, the relationship between geomagnetic activity and the stock market is still found further verifies our primary assumption. Because of the high correlation between YEAR and ILLIQ, we also do robust tests to rule out the multicollinear problem ( $\kappa=370.25$ ,  $VIF=2.03$ ).

In conclusion, we explore our research questions by examining monthly geomagnetic data (DGD) Ap indices and S&P stock market indices from 1998 to 2017. Our primary analysis tests whether the geomagnetic activity is related to the stock market. There is an inverse connection between geomagnetic activity and the stock market in the US. This increase is economically significant. The predicted probability of issuing an adverse stock signal is 1.2%. Furthermore, the relationship could be evolved into a causality. The Granger Causality Test is valid for both lag orders 1 and 2. Together, the results are consistent with previous research that geomagnetic activity and stock market are related indirectly or directly. One explanation for our findings is that annual or semiannual variations caused by geomagnetic activity affect investors' mood unconsciously or subconsciously, which leads to the waves in the stock market.

## 5. Conclusions

Using monthly-based geomagnetic indices and US stock market indices for the recent 20 years, we find compelling evidence supporting a causal relation between geomagnetic activity and stock return. First, supported by the results from the Granger Causality Test, geomagnetic activity is negatively related to the US stock market. Secondly, by regressing SAD on DGD and decomposing DGD, the seasonal variations of DGD on SMI are not directly supported, although the results represent the variations in winter and summer. At last, this research finds that market liquidity is positively related to the geomagnetic effect. To dig out more explanations for the influence of geomagnetic activity on the stock market, we view the potential relation between geomagnetic activity and behaviour finance as an avenue for future research.

This research offers several contributions. On the one hand, we introduce a geomagnetic relationship to the financial economics literature. Instead of dealing with extreme events such as geomagnetic storms, this research avoids the influence of virtual prediction on a strong result by exploring the causality effectiveness of monthly Ap index on the monthly US stock index. We view the potential relation between geomagnetic activity and behaviour finance as an avenue for future research. On the other hand, we contribute to existing research in financial economics examining whether geomagnetic activity pessimism impacts stock returns. Further, we explore how the effect of geomagnetic varies under two conditions which are (1) periodic effect of geomagnetic activities on the stock market and (2) liquidity effect on the correlation between Ap stock market indices. By focusing on the illiquidity and periodic output, this paper provides compelling support for the two conditions and reinforces existing studies.

This relationship and our findings should be of interest to the broader literature studying the geomagnetic activity on capital market effects of limited attention and also be of interest to study the geomagnetic activity effect on stock market behaviour. The geomagnetic effect on human investment behaviour will be discussed in future study.

## Data Availability

The geomagnetic index data and US stock market index data used to support the findings of this study are included in the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research was supported by the Natural Science Foundation of China (Grant# 71671154).

## References

- [1] H. Friedman, R. O. Becker, and C. H. Bachman, "Geomagnetic parameters and psychiatric hospital admissions," *Nature*, vol. 200, no. 4907, pp. 626–628, 1963.
- [2] M. N. Zhadin, "Review of Russian literature on biological action of DC and low-frequency AC magnetic fields," *Bioelectromagnetics: Journal of the Bioelectromagnetics Society*, vol. 22, no. 1, pp. 27–45, 2001.
- [3] S. J. Palmer, M. J. Rycroft, and M. Cermack, "Solar and geomagnetic activity, extremely low frequency magnetic and electric fields and human health at the Earth's surface," *Surveys in Geophysics*, vol. 27, no. 5, pp. 557–595, 2006.
- [4] B. Mendoza and S. S. de la Peña, "Solar activity and human health at middle and low geomagnetic latitudes in Central America," *Advances in Space Research*, vol. 46, no. 4, pp. 449–459, 2010.
- [5] I. Zakharov and O. Tyrnov, "The effect of solar activity on ill and healthy people under conditions of nervous and emotional stresses," *Advances in Space Research*, vol. 28, no. 4, pp. 685–690, 2001.
- [6] F. Perfetto, R. Tarquini, L. Tapparini, and B. Tarquini, "Influence of non-insulin-dependent diabetes mellitus on plasma endothelin-1 levels in patients with advanced atherosclerosis," *Journal of Diabetes and its Complications*, vol. 12, no. 4, pp. 187–192, 1998.
- [7] I. Stoilova and S. Dimitrova, "Geophysical variables and human health and behavior," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 70, no. 2-4, pp. 428–435, 2008.
- [8] J. S. Lerner, Y. Li, P. Valdesolo, and K. S. Kassam, "Emotion and decision making," *Annual Review of Psychology*, vol. 66, pp. 799–823, 2015.
- [9] H.-C. Shu, "Investor mood and financial markets," *Journal of Economic Behavior & Organization*, vol. 76, no. 2, pp. 267–282, 2010.
- [10] V. A. Belkin, "Magnetic solar and economic cycles: mechanism of close connection," *Economy of Region*, vol. 2013, no. 1, pp. 216–226, 2013.
- [11] A. Krivelyova and C. Robotti, *Playing the Field: Geomagnetic Storms and International Stock Markets*, Federal Reserve Bank of Atlanta, Working Paper No. 2003-5a, 2003.
- [12] M. M. Dowling and B. M. Lucey, "Are weather induced moods priced in global equity markets," *Journal of Multinational Financial Management*, vol. 18, no. 2, 2005.
- [13] M. Dowling and B. M. Lucey, "Mood and UK equity pricing," *Applied Economics Letters*, vol. 4, no. 4, pp. 233–240, 2008.
- [14] M. G. D. Feo, L. Juden-Kelly, S. Koren, and C. Whissell, "Geomagnetic variables and US market descriptors as predictors of the frequency of use of the Google search terms anxiety, depression, peace and dow jones," *Advances in Social Sciences Research Journal*, vol. 5, no. 8, pp. 517–521, 2018.
- [15] A. L. Cortie, "Sun-spots and terrestrial magnetic phenomena, 1898-1911," *Monthly Notices of the Royal Astronomical Society*, vol. 73, no. 1, pp. 52–60, 1912.
- [16] S. Chapman and J. Bartels, *Geomagnetism*, vol. 2, chapter 11, Clarendon Press, 1940.
- [17] M. J. Kamstra, L. A. Kramer, and M. D. Levi, "Winter blues: a SAD stock market cycle," *American Economic Review*, vol. 93, no. 1, pp. 324–343, 2003.
- [18] M. G. Lanfear, A. Lioui, and M. G. Siebert, "Market anomalies and disaster risk: Evidence from extreme weather events," *Journal of Financial Markets*, 2018.

- [19] R. F. Engle and C. W. J. Granger, "Co-integration and error correction: representation, estimation, and testing," *Econometrica*, vol. 55, no. 2, pp. 251–276, 1987.
- [20] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–238, 1969.
- [21] Y. Amihud, "Illiquidity and stock returns: cross-section and time-series effects," *Journal of Financial Markets*, vol. 5, no. 1, pp. 31–56, 2002.
- [22] H. O. Balli and B. E. Sørensen, "Interaction effects in econometrics," *Empirical Economics*, vol. 45, no. 1, pp. 583–603, 2013.
- [23] J. P. Forgas, "Mood effects on decision making strategies," *Australian Journal of Psychology*, vol. 41, no. 2, pp. 197–214, 1989.
- [24] R. W. Kay, "Geomagnetic storms: association with incidence of depression as measured by hospital admission," *The British Journal of Psychiatry*, vol. 164, no. 3, pp. 403–409, 1994.

## Research Article

# Single and Multiwavelength Detection of Coronal Dimming and Coronal Wave Using Faster R-CNN

Zongxia Xie and Chunyang Ji 

*College of Intelligence and Computing, Tianjin University, Tianjin 300350, China*

Correspondence should be addressed to Chunyang Ji; [chunyang\\_ji@tju.edu.cn](mailto:chunyang_ji@tju.edu.cn)

Received 12 April 2019; Accepted 23 June 2019; Published 8 July 2019

Guest Editor: Huaning Wang

Copyright © 2019 Zongxia Xie and Chunyang Ji. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Automatic detection of solar events, especially uncommon events such as coronal dimming (CD) and coronal wave (CW), is very important in solar physics research. The CD and CW are not only related to the detection of coronal mass ejections (CMEs) but also affect space weather. In this paper, we have studied methods for automatically detecting them. In addition, we have collected and processed a dataset that includes the solar images and event records, where the solar images come from the Atmospheric Imaging Assembly (AIA) of Solar Dynamics Observatory (SDO) and the event records come from Heliophysics Event Knowledgebase (HEK). Different from the methods used before, we introduce the idea of deep learning. We train single-wavelength and multiwavelength models based on Faster R-CNN. In terms of accuracy, the single-wavelength model performs better. The multiwavelength model has a better detection performance on multiple solar events than the single-wavelength model.

## 1. Introduction

Solar eruption is an event that produces numerous energy, releasing a large amount of radiation as ejecting plasma into the heliosphere as coronal mass ejections (CMEs). CMEs can cause geomagnetic storms, which in turn may affect the reliability of power systems. In addition, CMEs can also affect space weather [1, 2]. The detection of CMEs is of vital importance. However, CMEs are difficult to detect as a result of the complex characteristics. CMEs are associated with many solar phenomena. Coronal dimming (CD) is temporally and spatially consistent with CMEs [3] and dimming-associated CMEs have much higher speed than others [4]. Coronal waves (CWs) can be driven and shaped by the expanding flanks of CMEs [5]. Therefore, CD and CW become the main objects of detection.

As for CD, several algorithms were proposed to identify it. An algorithm for fully automatic detection of CD was discussed by Bewsher et al. The image was subjected to a differential operation to form a base-difference image, and then the pixel intensity was processed by threshold to find dimming regions [6]. Podladchikova et al. continued to study this algorithm and performed minimum and maximum pixel mapping to describe dimming regions

[7]. In addition, the statistical distribution of the running difference image is calculated to find significant disturbances, i.e., dimming regions. Different from the above method, Krista et al. used the original nondifference image and the corresponding magnetograms to detect the dimming regions [8].

For the detection of CW, Thompson et al. manually identified it [9] in 2009 with data on Solar and Heliospheric Observatory (SOHO) between 1997 and 1998. Then, Nitta et al. did a similar job and identified the CW [10] observed by the AIA on board the SDO spacecraft [11] in 2003. Manual identification is completely dependent on the user and this method is only suitable for small dataset. Podladchikova et al. used automatic dimming detection algorithm NEMO to detect CW based on SOHO [7]. Wills-Davey et al. proposed Huygens tracking technique [12]. This algorithm used percentage base-difference images to identify the pulse by finding the peak intensity line corresponding to the peak of the Gaussian cross-section. The semiautomated detection algorithm, Coronal Pulse Identification and Tracking Algorithm (CorPITA), was proposed by Long et al. [13]. CorPITA used intensity profile techniques to identify the propagation pulse, tracked its entire evolution, and then returned its kinematics estimate.

In 2010, NASA launched the SDO spacecraft, which orbited the Earth to capture full-disk images of the sun. AIA, an independent instrument of SDO, captures eight high-definition (4096×4096) full-disk images of different wavelengths every 10 to 12 seconds. SDO produces approximately 1.5 TB of data per day so that manual and traditional identification methods are no longer suitable for current conditions. In the case of large data, an automatic and rapid detection algorithm of CD and CW is very research-oriented. In 2017, Kucuk et al. used deep convolutional neural networks to classify solar events [14], including Sigmoid (SG) and Flare (FL). At the same time, they also used a deep learning model to detect solar events, such as Active Region (AR) [15].

In the field of computer vision, object detection is to locate the position of the object in an image and to identify the category of the object. This is a fundamental problem, but it is also very challenging because there are many interference factors in the image, such as the background clutter and perspective changes. In recent years, deep neural networks have become a research hotspot, and there are many high precision object detection based on deep learning, for example, Single Shot MultiBox Detector (SSD) [16], You Only Look Once (YOLO) [17], Region-based Fully Convolutional Networks (R-FCN) [18], and Faster Regions with Convolutional Neural Networks Features (Faster R-CNN) [19]. Both SSD and YOLO are fully end-to-end detection models, but SSD has an additional layer to predict the default box offset to improve accuracy. Unlike them, R-FCN and Faster R-CNN use the Region Proposal Network (RPN) [18]. The RPN is a separate network that generates proposed regions and flags, which contain the object or not. The difference between R-FCN and Faster R-CNN is the bounding box classification. Classification feature used by the R-FCN is to crop from the last layer of base network.

In this paper, deep convolutional neural networks (Faster R-CNN) is introduced to detect CD and CW. The rest of the paper is as follows. Section 2 introduces the dataset. Section 3 states the deep learning algorithm. The experiments and results are discussed in Section 4. Section 5 gives the conclusion.

## 2. Data Preparation

In this section, we introduce our dataset, including data collection and processing.

**2.1. Data Collection.** The dataset consists of two parts. One is the solar images. The other is the solar event records, especially the location coordinate and category of the solar events. As for the images, they come from AIA module of SDO. AIA can capture high-definition images at multiple wavelengths, but the metadata of AIA are FITS format and cannot be used directly to train the detection model. We can process the metadata into JPEG (4096×4096) format to continue the work. For the solar event records, they are determined according to HEK that receives metadata from many automated event detection modules. The metadata contains a lot of information, including the location coordinates, the start time, the end time, and the category of the solar events.

In order to collect required data, we use the SunPy (<https://sunpy.org/>), which is a package for python language. Through the API of this package, we can get solar event records from HEK based on the time and the category of solar event. For example, we can search CD and CW from 2014 to 2017, and then we can get CD and CW event records from different wavelength (171Å, 193Å, 211Å). From the event record, we can obtain the start time and the end time. According to them and the API, the metadata of the corresponding event can be downloaded from AIA module of SDO. The metadata for the event is recorded every 12 seconds in the AIA module of SDO, and we can get multiple metadata during the time of the event. In addition, we use the Solarsoft (<http://www.lmsal.com/solarsoft/>), which is a suite for IDL, to parse and preprocess metadata. What is more, the metadata are converted in turn into the JPEG image format. The event record also contains the location coordinates, the lower left and upper right coordinate of the bounding box. But these acquired coordinates are based on the Helioprojective Cartesian (HPC) coordinate system and need to be converted to the image pixel coordinates. We can do this using the API provided by the suite. So far, we have completed the data collection work, and then we process and clean the data.

**2.2. Data Processing.** There are many data sources in the HEK database, and the data is relatively messy. Inevitably there are some incorrect data. Furthermore, the occurrence of a solar event is a dynamic process, but the HEK database only provides a set of coordinates, which is not reasonable. We need to clean and process the data to make it more reliable. As mentioned in the description of the data collection above, we can get multiple metadata in the time range of the event. Then, we use the API to differentiate these metadata to get the running difference images. It helps us find the location of the event more easily. At the same time, we compare the location coordinates in the event record, clean up some unreasonable data, and simply adjust the location coordinates of the event. Among them, running difference image is a work that takes a lot of time, and it is necessary to constantly adjust some parameters to make the event noticeable. Figures 1 and 2 show some examples of CD and CW in the dataset, respectively.

In the end, we divide the dataset into eight parts. As for CD, it consists of four parts, three of which are single-wavelength (171Å, 193Å, 211Å) and one of which is multi-wavelength. Each image in the dataset has 4096×4096 pixels. Each part consists of a training set and a test set. It is worth mentioning that we collect data from 2014 to 2017, but the training and test sets are randomly divided. The three single-wavelength training sets and test sets are different. The multiwavelength training set is a union of single-wavelength training sets, and the test set is consistent with a single-wavelength test set. The remaining four parts are CW, and the situation is consistent with CD. In Table 1, we give details of the training set and test set. As we all know, CD and CW are not common solar events, especially CW. Such events are still controversial, and detecting them is a very challenging task.

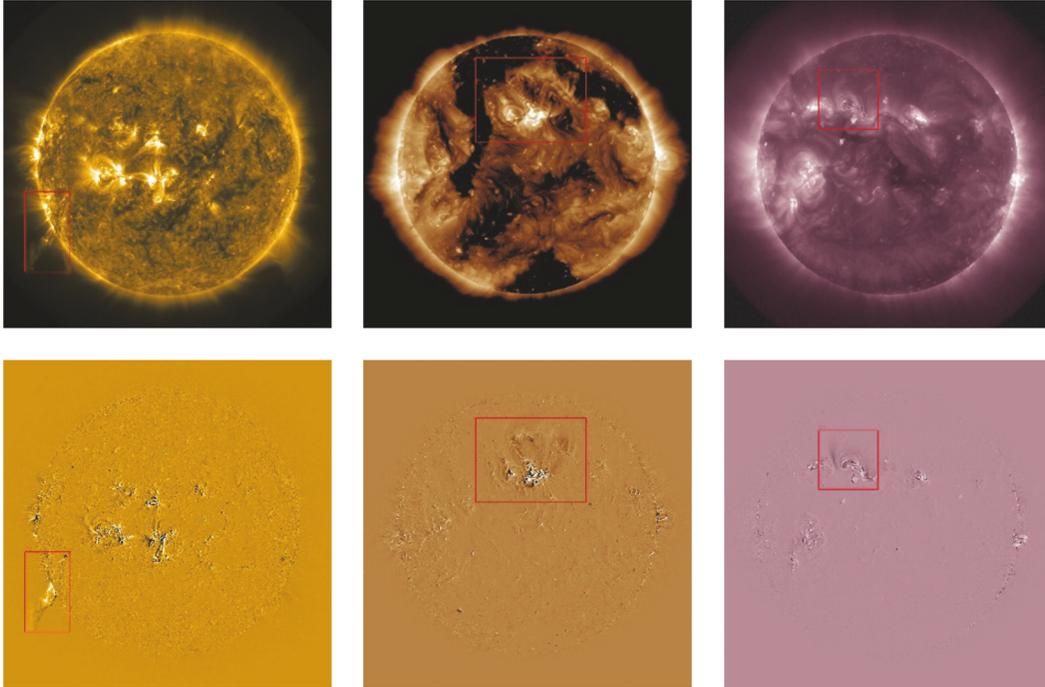


FIGURE 1: CD label. From left to right, the wavelength of 171Å, 193Å, and 211Å. Training images (top) and run-difference images (bottom).

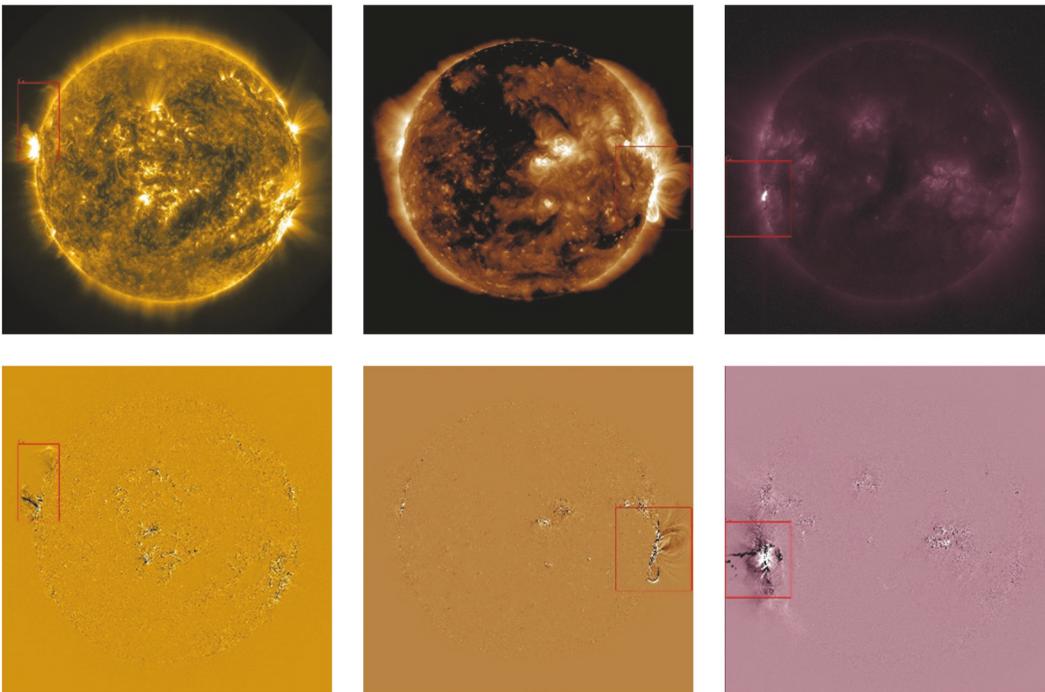


FIGURE 2: CW label. From left to right, the wavelength of 171Å, 193Å, and 211Å. Training images (top) and run-difference images (bottom).

### 3. Deep Learning for CD and CW Detection

In this section, we introduce the deep learning model, Faster R-CNN, for detecting CD and CW.

*3.1. Deep Learning and Convolutional Neural Networks.* Traditional machine learning techniques are difficult to perform object detection task. Deep learning is a relatively new field of research in machine learning. It is also known for its

TABLE 1: Training and test sets of single-wavelength and multi-wavelength for CD and CW. CD/CW-171, 193, 211 represents the model where the images of single-wavelength are used to train/test, respectively while CD/CW-M represents the model where multiple-wavelength images are used for training or testing.

Number	Event-wavelength	Training	Test
1	CD-171	224	71
2	CD-193	210	66
3	CD-211	187	24
4	CD-M	224+210+187=621	71+66+24=161
5	CW-171	363	114
6	CW-193	627	384
7	CW-211	374	47
8	CW-M	363+627+374=1364	114+384+47=545

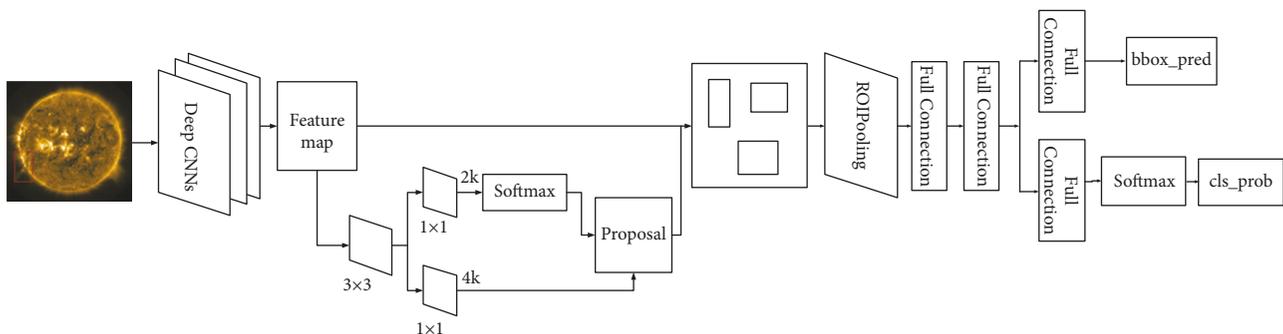


FIGURE 3: Network structure for detecting solar events using Faster R-CNN.

strong learning ability. The essence of deep learning is that features are not artificially extracted and specified, but are characterized by data learning.

The convolutional neural networks (CNNs) are a kind of deep learning model. By training the filters, the convolution and the pooling are alternately operated on the original input image, thereby the increasingly complex features of the hierarchy are extracted. CNNs can build a map between input and output, which is nonlinear. Compared to traditional algorithms, deep CNNs perform well in the face of big data. Kucuk et al. used deep convolutional neural networks to classify solar events, and the model can classify them correctly with high precision.

**3.2. Faster R-CNN.** Faster R-CNN is a faster object detector based on the deep learning framework named Caffe [20]. Before introducing Faster R-CNN, we briefly describe fast region-based convolutional neural networks (Fast R-CNN).

The main idea of Fast R-CNN is to turn the object detection problem into a classification problem based on candidate feature extraction. This allows taking advantage of strong performance of deep learning on classification issues. An image is input to the Fast R-CNN, and CNNs are first used to obtain the feature of the image. CNNs are called pretraining networks. Fast R-CNN uses Selective Search (SS), which is an algorithm to generate the object candidate regions [21]. At the same time, it proposes the Regions of Interest (RoI), which maps the candidate regions generated by the SS to the feature layers of the CNNs and extracts the

deep features directly on the feature layer. Then, softmax is used to classify the extracted features, and the regression of bounding box is trained to improve the accuracy of the object location. It can be seen that Fast R-CNN integrates the training, classification, and regression of CNNs to improve the efficiency of object detecting as a whole. But the SS algorithm is independent of the deep neural networks and cannot use GPU computing. The calculation speed has a certain impact on the overall performance of the algorithm. In order to solve this problem, Faster R-CNN was proposed.

Faster R-CNN is improved based on the Fast R-CNN algorithm. Instead of the SS algorithm in Fast R-CNN, Region Proposal Networks (RPN) is proposed for the purpose of extracting candidate regions. Faster R-CNN integrates candidate region extraction, deep feature extraction, classification, and bounding box regression into deep neural networks. All of these tasks can be trained in the GPU, increasing the efficiency of the algorithm without losing accuracy.

The input of the Faster R-CNN is the solar image and the corresponding label. The label includes the event type (CD) and location coordinates in the image. The feature is extracted from the solar image through a deep convolutional neural network, and then the feature map is processed in two lines. One line is a RPN network, and the purpose is to obtain candidate regions. RoI is operated on the obtained candidate regions and the feature map to get candidate region feature. The other line is the classification and bounding box regression of candidate regions feature. The topology of the network is shown in Figure 3.

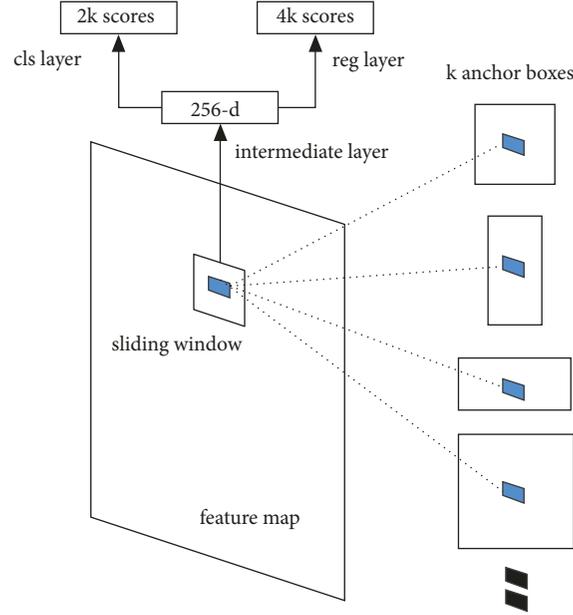


FIGURE 4: Network structure for RPN.

RPN is a full convolutional neural network. The previous layer of the input of the network is feature map, and the output is rectangle candidate region. First, a sliding window is performed on the feature map of the shared convolution network to obtain the  $n \times n$  window. This window is used as an input to the network, and  $k$  candidate regions are predicted at the same time. Here, we name the candidate region an anchor and each anchor has different sizes and ratios. Each window is mapped to a low latitude vector by a convolution operation. This vector is used in two subnetworks, with bounding box regression and bounding box classification. The output of the bounding box classification network is whether each anchor contains an object, i.e., foreground or background, and the size of the output vector is  $2k$ . The output of the bounding box regression network is to adjust the location of each anchor to make it more accurate, and the size of the output vector is  $4k$ . The network structure of RPN is shown in Figure 4.

After the RPN network, we get the object candidate region, which is also the rough position of the object. The deep feature is extracted by performing a RoI operation according to the object candidate region and the feature map. The deep feature is subjected to two full connection layer operations, and then the feature of the full connection is operated in parallel. One is the softmax classification layer, which outputs the probability distribution of each candidate region feature in the category while the other is the bounding box regression layer, which again corrects the location of the object to make it more accurate. The loss function has a combination of these two parts as

$$L(p, u, t^u, v) = L_{cls}(p, u) + \lambda [u \geq 1] \cdot L_{loc}(t^u, v) \quad (1)$$

where  $p = (p_0, \dots, p_k)$  represents the probability distribution of each candidate region feature on  $k + 1$  categories,  $u$  represents the real category,  $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$  is the output

of the bounding box regression, and  $v$  represents the ground truth of the category  $u$ .  $L_{cls}$  is the logistic loss of the real category  $u$  as

$$L_{cls}(p, u) = -\log(p_u) \quad (2)$$

$L_{loc}$  is the bounding box regression loss as (3). It is based on two sets of parameters,  $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$ ,  $v = (v_x, v_y, v_w, v_h)$ .  $L_1$  loss is given in (4).

$$L_{loc}(t^u, v) = \sum_{i \in \{x, y, w, h\}} \text{Smooth}_{L_1}(t_i^u - v_i) \quad (3)$$

$$\text{Smooth}_{L_1}(x) = \begin{cases} 0.5x^2 & |x| < 0 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (4)$$

For  $[u \geq 1]$ , the value is 1 when  $u \geq 1$ , indicating that the candidate region is positive and has regression loss; otherwise the value is 0, indicating no loss. At this time, the candidate region is the background and the background has no ground truth.  $\lambda$  is used to balance the two losses.

## 4. Experimental Results and Discussion

**4.1. Experiment Settings.** The experiment was performed on a GPU server, and the GPU type was NVIDIA GM200. In addition, we used the deep learning framework named Caffe. We used the datasets mentioned in the second section, including the training set and test set. For CD, because of small size of dataset, we used the ZF pretraining model. As for CW, the size is larger than CD, we use the VGG1024 pretraining model. We trained two different models, single-wavelength and multiwavelength and also tested the model with the same test set. All models were iterated 35000 times. It takes a long time to train a model, which is about 18 hours.

TABLE 2: The Ap value of three wavelengths in single-wavelength and multi-wavelength using Faster R-CNN.

Model	171 Å	193 Å	211 Å
Single-wavelength(CD)	0.895	0.774	0.909
Multi-wavelength(CD)	0.889	0.826	0.909
Single-wavelength(CW)	0.786	0.754	0.783
Multi-wavelength(CW)	0.767	0.729	0.590

TABLE 3: The Ap value of three wavelengths in single-wavelength and multi-wavelength using R-FCN.

Model	171 Å	193 Å	211 Å
Single-wavelength(CD)	0.864	0.758	0.887
Multi-wavelength(CD)	0.842	0.747	0.893
Single-wavelength(CW)	0.771	0.767	0.763
Multi-wavelength(CW)	0.764	0.749	0.638

In order to get the optimal performance, we only try several values of parameters, ‘iteration’ and ‘learning rate’.

We use average precision (Ap), precision, and recall to characterize the performance of the model. In order to test the model, it is very important to have a correct test standard. If the value (the ratio of intersection to union, IOU) between predicted bounding box (PBbox) and the ground truth box (GTbox) is greater than 0.5, the detection is considered to be correct. Recall is the ratio of corrected PBboxes in all ground truth as shown in (5). Precision is the ratio of corrected PBboxes in all PBboxes as shown in (6).

recall

$$= \frac{\text{number}(PBbox \mid IOU(PBbox, GTbox) \geq 0.5)}{\text{number}(GTbox)} \quad (5)$$

precision

$$= \frac{\text{number}(PBbox \mid IOU(PBbox, GTbox) \geq 0.5)}{\text{number}(PBbox)} \quad (6)$$

The Ap is related to the precision-recall curve and is calculated by the mean precision at eleven recall levels  $[0, 0.1, \dots, 1]$ :

$$Ap = \frac{1}{11} \sum_{r \in \{0, 0.1, \dots, 1\}} Pt(r) \quad (7)$$

Based on the level  $r$ , find the maximum precision value for which the corresponding recall exceeds  $r$  as

$$Pt(r) = \max_{r^*: r^* \geq r} p(r^*) \quad (8)$$

where  $P(r^*)$  is the precision at recall level  $r^*$ .

**4.2. Single-Wavelength Faster R-CNN.** For CD, the Ap value of 211 Å is 0.909, which is the best of three single-wavelength models. In contrast, the model of 193 Å has the lowest Ap value. But overall, the accuracy of the three single-wavelength models is very high, as shown in Table 2. In addition, we compare the predicted bounding box with the acquired labels

on the HEK database. The three models can accurately locate the location of the event. Some sample images of the detection results are shown in Figure 5.

As for CW, the accuracy of the single-wavelength detection models is approximately 78%. The Ap values of the 171 Å, 193 Å, and 211 Å models are shown in Table 2. Among them, the Ap value of 171 Å is the highest, that of 211 Å is the lowest, but the differences between the three models are not large. We also compare the test results of the models with the HEK database. The visual results are shown in Figure 5. We can see that the events marked in the HEK are recognized in our detection model.

In addition, we trained the R-FCN model to detect CDs and CWs. To ensure the effectiveness of the comparative experiments, we used the same dataset and experimental environment. In Table 3, as for CW, the Ap of the 193 Å is higher than the Faster R-CNN, and the Ap of multiwavelength detection model is higher than the Faster R-CNN at 211 Å. The other cases are not as good as the Faster R-CNN. We also recorded the training time of the two models. For the single-wavelength model, Faster R-CNN takes approximately 18 hours and R-FCN takes 20 hours. For the multiwavelength models, Faster R-CNN takes 21 hours and R-FCN takes 23 hours. Considering both factors, Faster R-CNN is better than R-FCN.

**4.3. Multiwavelength Faster R-CNN.** Since the single-wavelength model has good accuracy and intuitive effects, we mix the characteristics of the three wavelength solar events and complement each other to form a multiwavelength detection model. As for CD and CW, the Ap values of the detection model are also shown in Table 2. Compared with the results of the single-wavelength detection model, only the test images of 193 Å have higher Ap than single-wavelength models, while those of others have lower Ap than single-wavelength models. But this does not mean that the effect of multiwavelength is not as good as single-wavelength. Precision and recall are also indicators of model performance. The precision and recall of the events for single-wavelength and multiwavelength are shown in Figure 6. Here, the precision and recall is the mean of those on each test image. We find that the precision

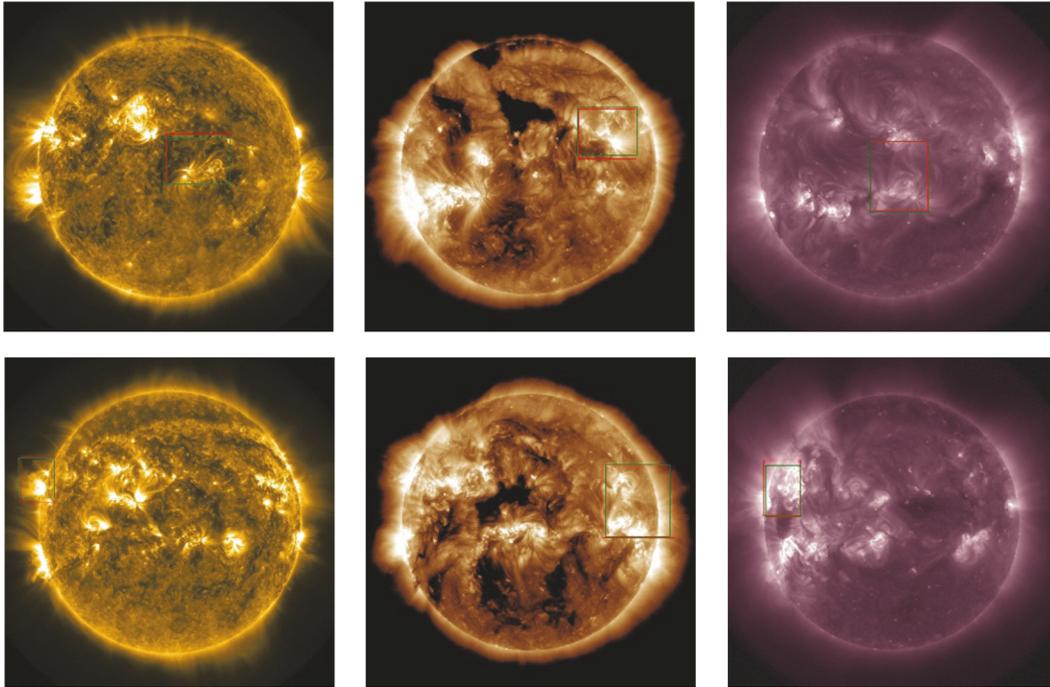


FIGURE 5: CD (top) and CW (bottom) test results. The red rectangle represents ground truth and the green rectangle represents the predicted bounding box of the single-wavelength model using Faster R-CNN.

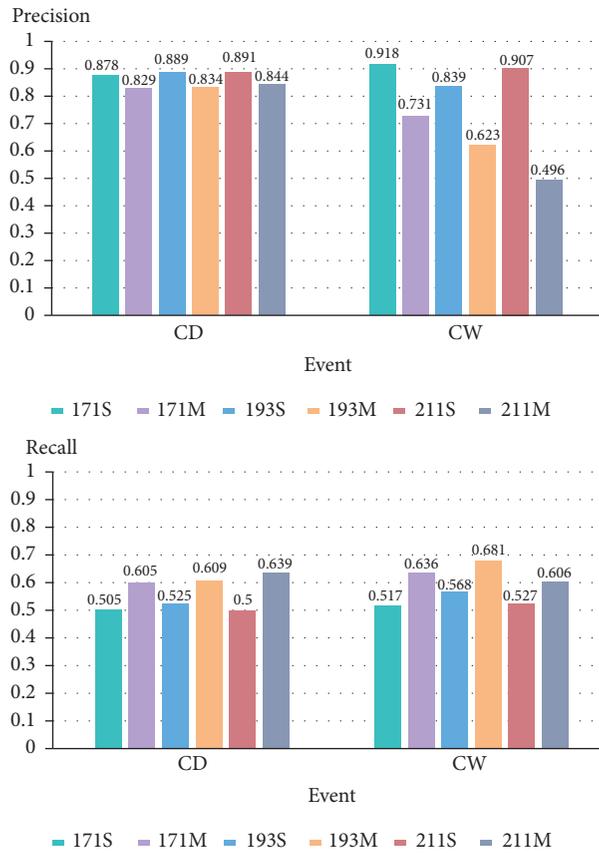


FIGURE 6: Precision and recall for different models. S and M represent the models where single-wavelength and multiwavelength images are used to train/test, respectively.

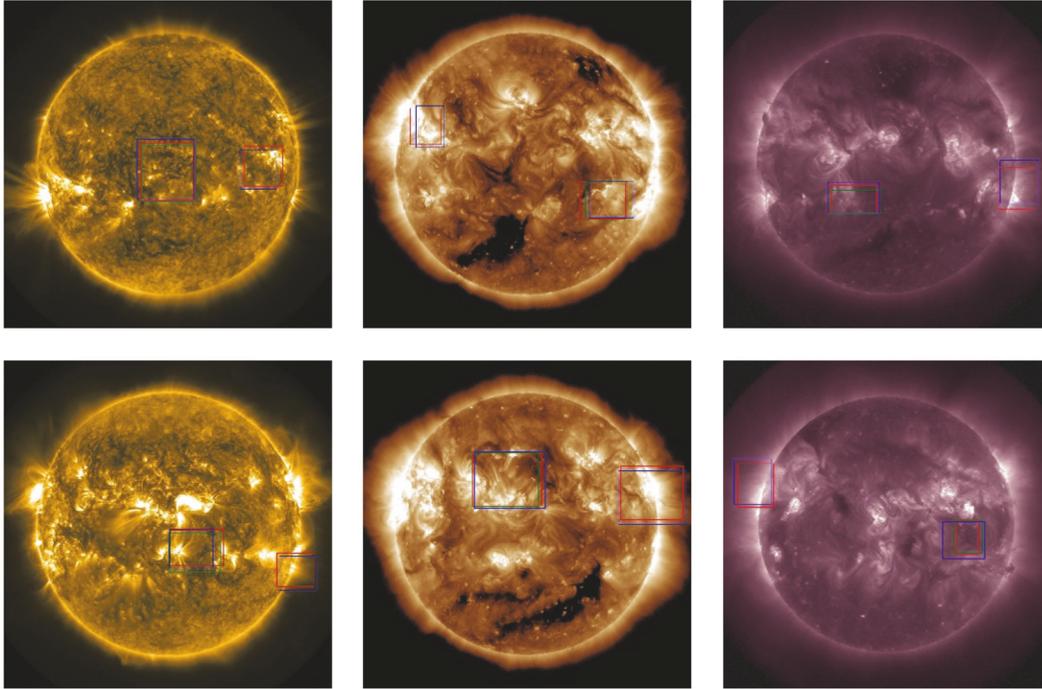


FIGURE 7: CD (top) and CW (bottom) test results. The red rectangle represents ground truth and the green rectangle represents the predicted bounding box of the single-wavelength model. The blue rectangle represents the predicted bounding box of the multiwavelength model using Faster R-CNN.

of the single-wavelength model is higher than that of the multiwavelength one, but the recall of the multiwavelength model is higher than that of the single-wavelength one. We can visually check the detection effect of multiple models. Compared to the single-wavelength model, we can see that the multiwavelength models are better for detecting multiple solar events. As shown in Figure 7, the original multiple solar events are not completely detected with single-wavelength, but the multiwavelength model is relatively accurately detecting those events. However, there are cases where the multiwavelength model is used to detect some incorrect bounding boxes. Because of these incorrect bounding boxes, it will lead to a decrease in the accuracy rate. Therefore, the precision of the multiwavelength model is lower than that of the single-wavelength model. For the recall rate, since the number of Gtbox is fixed, the denominator in (5) is constant, and the number of correct bounding box detected increases, which results in an increase in the recall rate. Although  $A_p$  is an important indicator of the measurement model, it is not absolute. It can be seen from the above analysis that the multiwavelength detection model still has a certain effect.

## 5. Conclusions

We train the deep learning models to detect CD and CW for single-wavelength and multiwavelength, respectively. The results show that the models can accurately detect about 70% of CD and CW. Single-wavelength models have high

accuracy in detecting single solar events, but sometimes it is not comprehensive when they detect multiple solar events. Compared to single-wavelength models, multiwavelength models have better performance in detecting multiple solar events, but not as accurate as single-wavelength models. In addition, multiwavelength models predict some incorrect bounding boxes, which may misinterpret some regions as solar events. It is feasible to apply deep learning to the field of solar event detection and has certain research value.

For future work, we first consider adding difference images during the training of the model. In other words, the solar events in the difference image are easier to identify and may enhance the detection of the original image. Furthermore, we consider feature fusion of multiple wavelengths. The same solar event may appear in images of multiple wavelengths, and the feature of merging multiple wavelengths may enhance the detection of solar events at a certain wavelength.

## Data Availability

SolarData consists of two parts, one is CD and the other is CW. The CD consists of three wavelengths (171 Å, 193 Å, 211 Å). Each wavelength contains the solar images and labels (event type and coordinates). The situation of CW is consistent with CD. The SolarData data used to support the findings of this study have been deposited in the GitHub repository (<https://github.com/Gkllmon/SolarData.git>).

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgments

The authors thank SunPy developers, SDO, and HEK for providing the data used in the paper. This work is supported by the National Natural Science Foundation of China under Grants 61432011, U1435212, 61105054, and Open Research Program of Key Laboratory of Solar Activity.

## References

- [1] N. Gopalswamy, "A global picture of CMEs in the inner heliosphere," in *The Sun and the Heliosphere as an Integrated System*, vol. 317 of *Astrophysics and Space Science Library*, pp. 201–251, Springer, Dordrecht, Netherlands, 2004.
- [2] I. G. Richardson, E. W. Cliver, and H. V. Cane, "Sources of geomagnetic storms for solar minimum and maximum conditions during 1972–2000," *Geophysical Research Letters*, vol. 28, no. 13, pp. 2569–2572, 2001.
- [3] T. H. Howard and R. A. Harrison, "On the coronal mass ejection onset and coronal dimming," *Solar Physics*, vol. 219, no. 2, pp. 315–342, 2004.
- [4] A. A. Reinard and D. A. Biesecker, "The relationship between coronal dimming and coronal mass ejection properties," *The Astrophysical Journal*, vol. 705, no. 1, pp. 914–919, 2009.
- [5] M. Temmer, A. M. Veronig, N. Gopalswamy, and S. Yashiro, "Relation between the 3D-geometry of the coronal wave and associated CME during the 26 April 2008 event," *Solar Physics*, vol. 273, no. 2, pp. 421–432, 2011.
- [6] D. Bewsher, R. A. Harrison, and D. S. Brown, "The relationship between EUV dimming and coronal mass ejections," *Astronomy & Astrophysics*, vol. 478, no. 3, pp. 897–906, 2008.
- [7] O. Podladchikova and D. Berghmans, "Automated detection Of eit waves and dimmings," *Solar Physics*, vol. 228, no. 1–2, pp. 265–284, 2005.
- [8] L. D. Krista, A. Reinard, and A. Alysha, "Study of the recurring dimming region detected at AR 11305 using the COronal DIMming Tracker (CoDiT)," *The Astrophysical Journal*, vol. 762, no. 2, p. 91, 2013.
- [9] B. J. Thompson and D. C. Myers, "A catalog of coronal "EIT wave" transients," *The Astrophysical Journal Supplement Series*, vol. 183, no. 2, pp. 225–243, 2009.
- [10] N. V. Nitta, C. J. Schrijver, J. Carolus et al., "Large-scale coronal propagating fronts in solar eruptions as observed by the atmospheric imaging assembly on board the solar dynamics observatory - An ensemble study," *The Astrophysical Journal*, vol. 776, no. 1, pp. 1567–1579, 2013.
- [11] P. C. Martens, G. D. Attrill, A. R. Davey et al., "Computer vision for the solar dynamics observatory (SDO)," *Solar Physics*, vol. 275, no. 1–2, pp. 79–113, 2012.
- [12] M. J. Wills-Davey, "Tracking large-scale propagating coronal wave FRONTS (EIT waves) using automated methods," *The Astrophysical Journal*, vol. 645, no. 1, pp. 757–765, 2006.
- [13] D. M. Long, D. S. Bloomfield, P. T. Gallagher, and D. Pérez-Suárez, "CorPITA: an automated algorithm for the identification and analysis of coronal "EIT Waves"," *Solar Physics*, vol. 289, no. 9, pp. 3279–3295, 2004.
- [14] A. Kucuk, J. M. Banda, and R. A. Angryk, "Solar event classification using deep convolutional neural networks," in *Proceedings of the International Conference on Artificial Intelligence & Soft Computing*, Springer, Cham, Switzerland, 2017.
- [15] K. Ahmet, B. Aydin, and R. Angryk, "Multi-wavelength solar event detection using faster R-CNN," in *Proceedings of the 5th IEEE International Conference on Big Data, Big Data 2017*, pp. 2552–2558, December 2017.
- [16] W. Liu, D. Anguelov, D. Erhan et al., *SSD: Single Shot Multibox Detector*, vol. 9905, European conference on computer vision, Springer, Cham, Switzerland, 2016.
- [17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, July 2016.
- [18] J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: Object detection via region-based fully convolutional networks," in *Proceedings of the 30th Annual Conference on Neural Information Processing Systems, NIPS 2016*, December 2016.
- [19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 6, pp. 1137–1149, 2015.
- [20] Y. Jia, E. Shelhamer, J. Donahue et al., *Caffe: Convolutional Architecture for Fast Feature Embedding*, ACM, 2014.
- [21] J. R. R. Uijlings, K. E. A. Van De Sande, T. Gevers, and A. W. M. Smeulders, "Selective search for object recognition," *International Journal of Computer Vision*, vol. 104, no. 2, pp. 154–171, 2013.

## Research Article

# The Mid-Term Forecast Method of $F_{10.7}$ Based on Extreme Ultraviolet Images

L. Lei <sup>1,2</sup>, Q. Zhong <sup>1,2</sup>, J. Wang,<sup>2</sup> L. Shi,<sup>1,2</sup> and S. Liu<sup>1,2</sup>

<sup>1</sup>National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China

<sup>2</sup>University of Chinese Academy of Sciences, Beijing 100190, China

Correspondence should be addressed to Q. Zhong; zhongqz@nssc.ac.cn

Received 15 January 2019; Revised 22 March 2019; Accepted 16 April 2019; Published 13 May 2019

Academic Editor: Michael Kueppers

Copyright © 2019 L. Lei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The solar radio flux at 10.7cm ( $F_{10.7}$ ) is a direct monitor and an important indicator of solar variability, and  $F_{10.7}$  is commonly used in empirical atmospheric models, ionosphere models, etc. The source regions of  $F_{10.7}$  are mainly in the corona above the active regions, and the extreme ultraviolet (EUV) images reflect the coronal thermal structure. In this paper, an index is defined as  $P_{SR}$  based on the intensity values of solar EUV images to represent the coronal contribution to  $F_{10.7}$ . The Spearman correlation coefficient between the observed values of  $F_{10.7}$  and  $P_{SR}$  is 0.85 in 304 Å EUV images. Based on the high correlation, an empirical model is constructed. Combining the EUV data of SDO/AIA and the twin STEREO/EUVI, solar full-disk EUV images can be generated, and the future 27-day values of  $P_{SR}$  can be calculated. Then, a realistic estimation of  $F_{10.7}$  from 1 to 27 days in advance can be provided by the empirical model. Compared to the predictive values of  $F_{10.7}$  by the 54th-order autoregressive models in 2012-2013, the error drop-rate of our model is 12.54%, and our method has significant advantages in the upcoming 3 to 27 days' forecast.

## 1. Introduction

$F_{10.7}$  is an index of solar radio flux at the frequency of 2800 MHz and the wavelength of 10.7 cm. There are primarily two kinds of radiation mechanisms: gyroresonance emission and bremsstrahlung emission [1–4]. Gyroresonance emission only occurs where the magnetic field is sufficiently strong. Schonfeld et al. [5] found that, on 2011 December 9, in the rising phase of Cycle 24,  $8.1\% \pm 0.5\%$  of the variable component of the  $F_{10.7}$  flux was associated with the gyroresonance emission mechanism, although this percentage varies significantly over the activity cycle. The strength of bremsstrahlung emission is related to the plasma density. Because the active regions and the flares are denser than the quiet Sun, the source regions of  $F_{10.7}$  are mainly in the coronal heights above the active regions [6]. Livingston et al. [7] found that there is no more the linear relation traditionally assumed to exist between  $F_{10.7}$  and the sunspot number during Cycle 24. Selhorst et al. [8] found that  $F_{10.7}$  and the number of active regions correlate well during the period of 1992–2013. It may be caused that  $F_{10.7}$  are more sensitive to weaker magnetic fields than those necessary to form sunspots, of the order of

1500 G. The observation of  $F_{10.7}$  started in 1947, and it has never been interrupted. At present, the commonly used value of  $F_{10.7}$  in the world is observed at local noontime (20:00 UT) by the Dominion Radio Astrophysical Observatory in Canada, and it is expressed in units of *sfu* (solar radiation flux density,  $1 \text{ sfu} = 10^{-22} \cdot \text{W} \cdot \text{m}^{-2} \cdot \text{Hz}^{-1}$ ) [9].

The solar extreme ultraviolet (EUV) irradiance heats and ionizes the Earth's atmosphere. Because the EUV irradiance cannot be observed from the ground,  $F_{10.7}$  has been used as a proxy to represent EUV [5]. The upper atmospheric models have the computer code which also use  $F_{10.7}$  as the proxy for solar irradiances [10]. The model of upper atmosphere is commonly used to calculate the atmospheric density to the orbital prediction of low Earth orbit satellites, so it is necessary to input the future  $F_{10.7}$  in the case of orbit prediction [11–13]. Therefore, the prediction of  $F_{10.7}$  plays a major role in the accuracy of orbit prediction.

The approximately 11-year long-period and average 27-day mid-period variations have already been discovered in solar activity. The 11-year periodicity is the solar variational cycle, and the 27-day periodicity is due to the period of the

solar rotation [14]. Prediction of the solar radiation index should correspond with the periodic features of solar activity. Singular spectrum analysis (SSA) is a nonparametric spectral estimation method, which can be an aid in the decomposition of time series into a sum of components and forecast the time series by these components. SSA can be effectively used to analyze the time series with periodic oscillations [15]. In machine learning, an autoregressive model (AR-model) could learn from a series of timed steps and take the previous results as inputs for a regression model, in order to predict the value of the next time step [16]. Thus, Zhong et al. [17] and Liu et al. [18] used the SSA method and the 54th-order AR-model, respectively, to forecast 27-day  $F_{10.7}$  because both the SSA method and the AR-model are suited to periodic and cyclical statistics. The methods of these two articles both showed predictive precision higher than that achieved by the United States Air Force (USAF) during the solar minimum in the Cycle 23. However, when a large active region (AR) rapidly appears on or disappears from the solar disk, the predictive precision of the 54th-order AR-model is unsatisfactory. Additionally, Wang et al. [19] proposed a mathematical method to extend the mid-term prediction of  $F_{10.7}$  to as long as 54 days without the need for extra solar observation data.

$F_{10.7}$  is generated at coronal heights and related to the presence of active regions and the occurrence of flares [8]. Some empirical prediction models based on major solar features are described below. Wen et al. [20] forecast  $F_{10.7}$  using the areas and positions of sunspots. Henney et al. [21] forecast  $F_{10.7}$  utilizing advanced predictions of the global solar magnetic field generated by a flux transport model. The two prediction models above were constructed only using data from the Earth-side solar hemisphere; many scholars have focused on the far-side solar disk. Quémerais and Bertaux [22] forecast future 14-day  $F_{10.7}$  using the interplanetary Lyman  $\alpha$  background data obtained by the Solar Wind Anisotropies (SWAN) telescopes on board the Solar and Heliospheric Observatory (SOHO) [23]. Lindsey and Braun [24] used seismic maps of the far-side solar hemispheres for solar active regions (ARs) forecasting. Although the Lyman  $\alpha$  background data and seismic maps of the far-side solar disk have the longer observed time span, the extreme ultraviolet (EUV) images of the twin Solar Terrestrial Relations Observatory (STEREO) spacecraft used in this study can show the far-side solar ARs more directly and describe their real-time variation characteristic more accurately.

The different wavelength EUV images are expected to observe the different regions of the solar atmosphere, such as the coronal hole (CH), quiet sun (QS), active region (AR), and flare plasma [25], and the values of  $F_{10.7}$  have a high correlativity with the magnetic active regions [26].  $F_{10.7}$  can be visualized as viewing the chromosphere (electron temperature  $\sim 10,000\text{K}$ ) through a lower corona of varying optical thickness due to the changing extent and density of the trapped coronal plasma overlying active regions and other active structures. For the discussion we formulate a new index,  $P_{\text{SR}}$ , which is a proxy for the low-coronal, free-free contribution to  $F_{10.7}$ , derived from the EUV data of the Atmospheric Imaging Assembly (AIA) on board

the Solar Dynamics Observatory (SDO) [27]. Balan et al. [28] demonstrated that the shorter (extreme ultraviolet and ultraviolet) and longer (10.7 cm) wavelength solar fluxes have a nonlinear relationship during entire solar cycles. This is because the decreasing presence of gyroresonant absorption compared with free-free emission at low activity levels and the contribution of the magnetic field strength below about 1500 Gauss to  $F_{10.7}$  is undercounted [29]. Therefore, we define a function of  $F_{10.7}$  in terms of  $P_{\text{SR}}$ . Furthermore, since February 2011, the entire far-side solar disk could be observed directly by the Extreme Ultraviolet Imager (EUVI) instruments, the telescopes in the two STEREO's Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI) imaging suite [30]. The solar full-disk EUV images can be obtained by combining data from STEREO/EUVI with that of SDO/AIA. Then, the future 27-day values of  $P_{\text{SR}}$  can be calculated to forecast  $F_{10.7}$ . In this paper, an empirical method is constructed according to the ideas above to predict 27-day  $F_{10.7}$  based on the EUV images.

The aim of this paper is to predict solar  $F_{10.7}$  up to one Carrington rotation in advance utilizing full solar composite 304 Å EUV images from SDO/AIA and STEREO/EUVI. We first defined an index as  $P_{\text{SR}}$  based on the intensity values of solar EUV images to represent the low-coronal, free-free contribution to  $F_{10.7}$ . This paper importantly demonstrates the tangible benefits that 360 degree solar observations provide for the prediction of solar activity. This prediction method is found to perform better in the 3-27 days period, particularly in cases where active regions emerge on the far side of the Sun.

After introducing the background of forecasting  $F_{10.7}$  in Section 1, Section 2 presents and analyzes the dataset and establishes the method. Section 3 contains the results and discussion, and Section 4 presents the conclusion.

## 2. Data Processing and Method Establishment

The two EUV channels (171 Å and 304 Å) of SDO/AIA are chosen, because they are also in STEREO/EUVI. The daily level-1  $512 \times 512$  data FITS (Flexible Image Transport System) files of SDO/AIA from May 2010 to December 2015 are downloaded from the available database of the Joint Science Operations Center (JSOC) at Stanford University (<http://jsoc.stanford.edu/>), and the daily  $512 \times 512$  data FITS files of STEREO/EUVI from January 2011 to December 2013 are downloaded from the available database of the website <https://stereoftp.nascom.nasa.gov/data/beacon/ahead/secchi/img/euvi/>.

The above data files are updated daily at approximately 20:00 UT, which corresponds to the observed time of the  $F_{10.7}$  index. The FITS files are eliminated if their quality (FITS header keyword) is not equal to 0. The data files of  $F_{10.7}$  can be downloaded from the websites of the National Oceanic and Atmospheric Administration (NOAA) ([ftp://ftp.ngdc.noaa.gov/STP/space-weather/solar-data/solar-features/solar-radio/noontime-flux/penticton/penticton\\_observed/listings/listing\\_drao\\_noontime-flux-observed\\_daily.txt](ftp://ftp.ngdc.noaa.gov/STP/space-weather/solar-data/solar-features/solar-radio/noontime-flux/penticton/penticton_observed/listings/listing_drao_noontime-flux-observed_daily.txt)).

The SDO/AIA datasets from May 2010 to December 2015 are used to determine the form of the function between  $F_{10.7}$  and  $P_{SR}$ . The SDO/AIA datasets of 2012-2013 and STEREO/EUVI datasets are taken as the testing sets.

**2.1. Coronal Contribution to  $F_{10.7}$ .** To match the EUVI data, the previous work should use the overlapping and interchangeable AIA and EUVI data to represent measurements from the same plasma [31]. This previous work has been developed as an automated procedure in the SSWIDL software routines (*ssc\_form\_euvi\_synoptic.pro*). In this routine, to match the SECCHI/EUVI data, the SDO/AIA data has been processed by the following formulas:

$$\text{For } 171\text{\AA}, \text{data}_{out} = \text{data}_{aia} / \text{exptime} \times 1.1$$

$$\text{For } 304\text{\AA}, \text{data}_{out} = \text{data}_{aia} / \text{exptime} \times \text{ssc\_get\_aia\_304\_factor}(\text{date}_{obs})$$

The “ $\text{data}_{aia}$ ” is the original SDO/AIA data; the “ $\text{exptime}$ ” is the exposure time of SDO/AIA in seconds. The “ $\text{date}_{obs}$ ” is date and time when observation of this image started. The parameter “1.1” is the rough conversion factor for AIA 171 Å images to EUVI 171 Å. The routine “*ssc\_get\_aia\_304\_factor.pro*” returns the rough conversion factor for AIA 304 Å images to EUVI 304 Å. Finally, the “ $\text{data}_{out}$ ” is the SDO/AIA data which matches the SECCHI/EUVI data. To calibrate the datasets of three EUV cameras, all SDO/AIA datasets have been processed by the above routine and all our results of SDO/AIA are based on the “ $\text{data}_{out}$ ”.

Vernazza et al. [32], Krista and Gallagher [33], and Pérez-Suárez et al. [34] found different solar regions (coronal holes, quiet sun, and active regions) corresponding to the different distributions in the intensity histogram of EUV images. Schonfeld et al. [5] showed that the EUV images collected by SDO/AIA could represent the bremsstrahlung component of  $F_{10.7}$ , so a proxy is defined in our paper to represent the coronal contribution to  $F_{10.7}$  in EUV images by the following formula:

$$P_{SR} = \sum_{i=1}^n I_i, \quad \text{if } I_i \geq I_{SR}. \quad (1)$$

where  $I_{SR}$  is the pixel intensity threshold of source regions (SRs). The pixel point belongs to an SR if  $I_i$  is greater than or equal to  $I_{SR}$ . The parameter  $n$  is the sum of pixel numbers on the Earth-side EUV images, so  $P_{SR}$  is the integral effect of all SRs in the Earth-side corona.

The Spearman Regression correlation coefficient ( $R_S$ ) is defined as the Pearson correlation coefficient between the ranked variables, and  $R_S$  assesses how well the relationship between two variables can be described using a monotonic function.  $R_S$  is computed from

$$R_S = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y} \quad (2)$$

where  $\text{cov}(x, y)$  is the covariance of the rank variables.  $\sigma_x$  and  $\sigma_y$  are the standard deviations of the rank variables [35, 36].

There is a nonlinear relationship between  $P_{SR}$  and  $F_{10.7}$  [28], so the  $R_S$  between them in different  $I_{SR}$  values is to define the values of  $I_{SR}$  in the EUV images of two channels. The

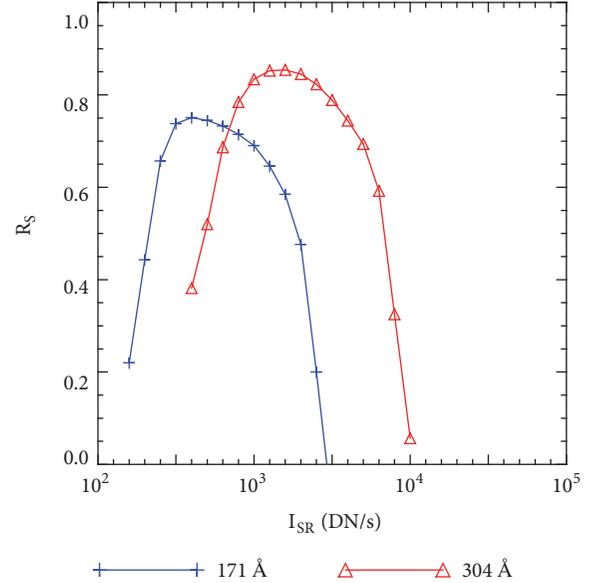


FIGURE 1: The Spearman correlation coefficients ( $R_S$ ) between  $F_{10.7}$  and  $P_{SR}$ , taking the specified pixel intensity as the threshold of the SRs in two channels’ EUV images from May 2010 to December 2015.

results are shown in Figure 1. The max  $R_S$  values are 0.75 in 171 Å and 0.85 in 304 Å.

The 171 Å passband mainly reflects the upper transition regions and the quiet coronal regions of the solar atmosphere [27]. The 304 Å passband shows the upper chromosphere/transition regions and less variability with solar activity on a short time scale [37].  $F_{10.7}$  can be visualized as viewing the chromosphere (electron temperature  $\sim 10,000\text{K}$ ) through a lower corona of varying optical thickness due to the changing extent and density of the trapped coronal plasma overlying active regions and other active structures [9]. The source regions of  $F_{10.7}$  are closer to the regions of solar 304 Å than 171 Å, so the agreement between the  $F_{10.7}$  and 304 Å is better than 171 Å. When  $I_{SR}$  is equal to  $10^{3.2}$  DN/s in 304 Å, the  $R_S$  between  $P_{SR}$  and  $F_{10.7}$  is up to 0.85. Thus, the pixels where intensities are greater than  $10^{3.2}$  DN/s show some information about the  $F_{10.7}$  source regions.

Based on the above analysis, the index associated with the  $F_{10.7}$  source regions in 304 Å EUV images is defined as follows:

$$P_{SR,304} = \sum_{i=1}^n I_i, \quad \text{if } I_i \geq 10^{3.2} \text{ DN/S}. \quad (3)$$

The green part in Figure 2(b) shows where the  $P_{SR,304}$  is on the EUV image. Most of this area covers the ARs, which is consistent with previous research: the source regions of  $F_{10.7}$  are mainly in the corona above the active regions [5, 6].

**2.2. Establishing the Method.** Figure 3 shows the scatter diagram between  $F_{10.7}$  and  $P_{SR,304}$  from May 2010 to December 2015. Considering the exponential relationship between

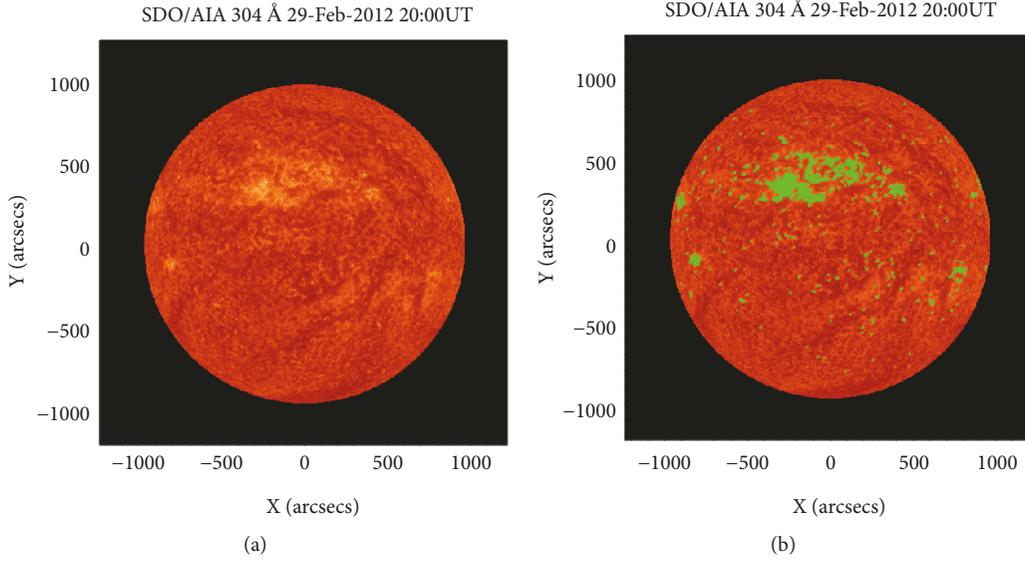


FIGURE 2: (a) The 304 Å EUV image on 29 February 2012; (b) pixels with intensities over  $10^{3.2}$  DN/s (green dots).

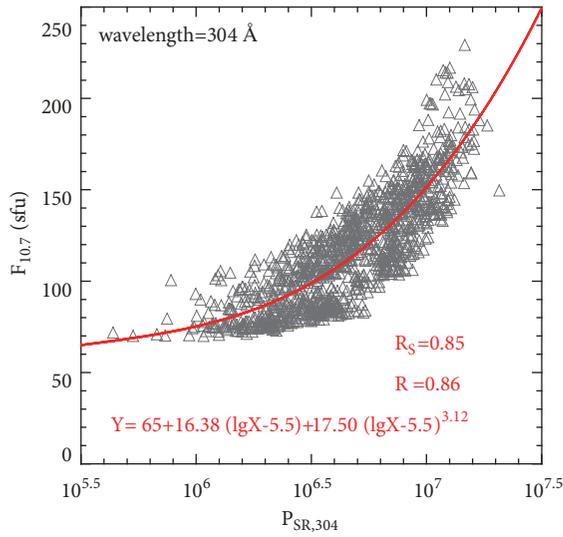


FIGURE 3: Scatter diagrams with the nonlinear function (solid red line) of  $P_{SR,304}$  and  $F_{10.7}$  from May 2010 to December 2015. The Spearman correlation coefficient ( $R_s$ ) is between  $P_{SR,304}$  and  $F_{10.7}$ , and the Pearson correlation coefficient ( $R$ ) is between fitted  $Y$  and  $F_{10.7}$ . The function of the solid red line is  $Y = 65 + 16.38(\lg X - 5.5) + 17.5(\lg X - 5.5)^{3.12}$ .

$P_{SR,304}$  and  $F_{10.7}$  in Figure 3, an empirical function of  $P_{SR,304}$  and  $F_{10.7}$  is defined as follows:

$$F_{10.7} = c + a_1 \cdot (\lg P_{SR,304} - 5.5) + a_2 \cdot (\lg P_{SR,304} - 5.5)^{a_3} \quad (4)$$

where  $c$  is a constant term that is expected to be the coronal base value. The parameters  $a_1$ ,  $a_2$ , and  $a_3$  are the undetermined coefficients.

To derive the best-fit coefficients for (4), all available daily  $P_{SR,304}$  and  $F_{10.7}$  from May 2012 to December 2015 are used for fitting by the nonlinear least square methods. The fitting function in Figure 3 is  $Y = 65 + 16.38(\lg X - 5.5) + 17.5(\lg X - 5.5)^{3.12}$ . The parameter “5.5” is the logarithmic total intensities while there is no active region on solar surface. That means the minimum value of  $F_{10.7}$  is equal to 65 *sfu* while  $\lg X$  equals 5.5. The higher order term means that the rate of  $F_{10.7}$  increasing with  $P_{SR,304}$  growth is nonlinear. To reduce the error of fitting function, we add a linear term. The correlation coefficient between fitted  $Y$  and  $F_{10.7}$  is 0.86. In consideration of the degradation of EUV instruments, the parameters  $a_1$ ,  $a_2$ , and  $a_3$  are calculated by sliding fitting the daily  $P_{SR,304}$  and  $F_{10.7}$  of previous 14 CRs (Carrington rotations, ICR = 27 days) using the nonlinear least square methods. The correlation coefficient of sliding fitting is up to 0.92 (in Figures 10(a) and 10(d)); these are described in detail on the part three.

The STEREO/EUVI system provides a direct observation of the far side of the solar disk, showing areas which will rotate onto the side of the Sun visible from Earth in the next few days. There is a special IDL procedure for generating the full-disk EUV image (Figure 4) in the SolarSoft tree for STEREO (`$$$SW/stereo/ssc/idl/beacon/ssc_form_euvi_synoptic.pro`) that combines the 304 Å data from SDO/AIA with the data from STEREO/EUVI nearest to 20:00 UT. The “+” symbols under the numbers 0-27 in Figure 4 are the diurnal projection of the Earth on the solar surface along the Earth-Sun line from 27 March 2013 to 27 days later. The latitude and longitude of Earth’s projection can be calculated by the IDL procedure in the SolarSoft tree for STEREO (`$$$SW/stereo/gen/idl/spice/get_stereo_lonlat.pro`). The green regions show where the pixel intensities are greater than  $10^{3.2}$  DN/s. The angle between the Earth-Sun line and the normal direction of pixel  $i$  in the EUV image is defined as  $\alpha_{i,t}$ . While the range of  $\alpha_{i,t}$  belongs to  $[0^\circ, 90^\circ]$ , the pixel  $i$  is in the Earth-side disk. Then converting into the same coordinates

TABLE 1: The coefficients ( $a_t$  and  $b_t$ ) and the correlation coefficient (R) of the fitting linear equation in Figure 5.

t (day)	$\lg(a_t)$	$b_t$	R	t (day)	$\lg(a_t)$	$b_t$	R
1	5.61	0.94	0.92	15	6.45	0.58	0.63
2	5.76	0.91	0.92	16	6.45	0.58	0.63
3	5.86	0.90	0.89	17	6.45	0.58	0.63
4	5.97	0.88	0.88	18	6.45	0.58	0.61
5	6.05	0.86	0.86	19	6.46	0.57	0.59
6	6.13	0.84	0.83	20	6.47	0.56	0.58
7	6.17	0.83	0.81	21	6.47	0.55	0.56
8	6.22	0.81	0.78	22	6.49	0.53	0.55
9	6.28	0.77	0.75	23	6.51	0.51	0.51
10	6.34	0.73	0.71	24	6.52	0.49	0.49
11	6.39	0.67	0.68	25	6.53	0.48	0.47
12	6.41	0.64	0.67	26	6.55	0.47	0.46
13	6.43	0.61	0.67	27	6.56	0.46	0.44
14	6.45	0.59	0.64				

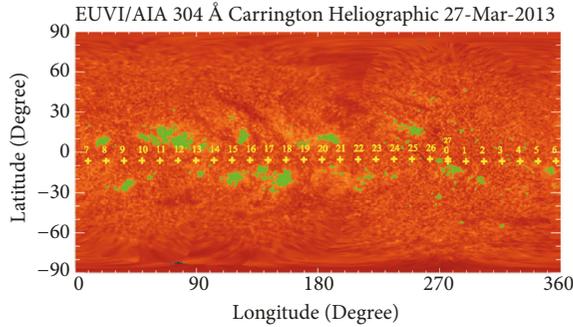


FIGURE 4: The 304 Å EUV image on 27 March 2013 from the combination of the 304 Å data from SDO/AIA with the data from the twin STEREO/EUVI. The coordinate is Carrington Heliography. The green regions represent the SRs of  $F_{10.7}$ , in which the pixel intensity is greater than or equal to  $10^{3.2}$  DN/s. The gold plus signs represent the projected position of Earth at 27 March 2013 (in day 0) and within 1-27 days later.

with SDO/AIA images, the equivalent predictive  $P_{SR,304}$  of the next 27 days from EUV images is defined as  $W_{304,t}$  in (5).

$$W_{304,t} = \sum_{i=1}^n I_{i,t}, \quad (5)$$

*if*  $I_{i,t} \geq 10^{3.2}$  DN/s,  $t = 1, 2, 3, \dots, 27$

where the parameter  $t$  represents the day. The parameter  $I_{i,t}$  is the intensity of pixel  $i$  in the next  $t$  days' image.

The area and intensity of AR will change in the next 27 days, especially in the solar maximum. The features of an AR's appearance, development, and disappearance are complex and unique. Therefore, the future values of  $W_{304,i}$  are very difficult to forecast even though the far-side solar disk is observed before 13.5 days. So we compare the previous  $P_{SR,304}$  with  $W_{304}$  on the same days to analyze the relationship between the  $P_{SR,304}$  and  $W_{304}$ . For example, the value of  $W_{304,1}$  on 27 March 2013 corresponds to the value of  $P_{SR,304}$  on 28

March 2013. The first scatter diagram at the top left corner of Figure 5 shows the linear relationship between the values of  $W_{304,1}$  from 2 January 2011 to 28 February 2013 and the values of  $P_{SR,304}$  from 3 January 2011 to 1 March 2013. The last scatter diagram at the bottom right corner of Figure 5 shows the relationship between the values of  $W_{304,27}$  from 2 January 2011 to 28 February 2013 and the values of  $P_{SR,304}$  from 29 January 2011 to 27 March 2013. Even though the points disperse and the linear correlation coefficients dwindle with the growing parameter  $t$ , we assumed there are 27 kinds of linear relationships between the values of  $W_{304,t}$  and  $P_{SR,304}$  in the next 27 days, which is defined in (6). After adjusting for (6), the parameter  $S_{304,t}$  is the corrected and predictive future 27-day values of  $P_{SR,304}$ .

$$S_{304,t} = a_t + b_t \cdot W_{304,t}, \quad t = 1, 2, 3, \dots, 27 \quad (6)$$

where  $a_t$  and  $b_t$  are the undetermined coefficients in the fitted linear equation (the green solid line in Figure 5), which are given in Table 1. The parameter  $t$  represents the day. Finally, the values of  $F_{10.7}$  can be predictive 27 days in advance through substitution of  $S_{304,t}$  for  $P_{SR,304}$  in (4).

### 3. Results and Discussion

In the application, there is no data after the forecasting date. Only the daily  $P_{SR,304}$  and  $F_{10.7}$  before the forecasting date can be used to obtain the best-fit coefficients for (4). The correlation between  $Y$  and  $F_{10.7}$  is performed for a certain number of Carrington rotations (CRs) before the start date of the series. This date is then advanced one day at a time through 2012 and 2013, generating a new correlation on each day. The minimum, average, and maximum correlation from this series is then recovered for that fitting window. The fitting window is varied from 1 to 28 CRs to generate Figure 6. The minimal correlation coefficient ( $R_{\min}$ ) reaches the maximum in 14 CRs, and the three parameters  $R_{\max}$ ,  $R_{\text{mean}}$ , and  $R_{\min}$  are steady after 14 CRs. Thus, the length of the fitting window is 14 CRs.

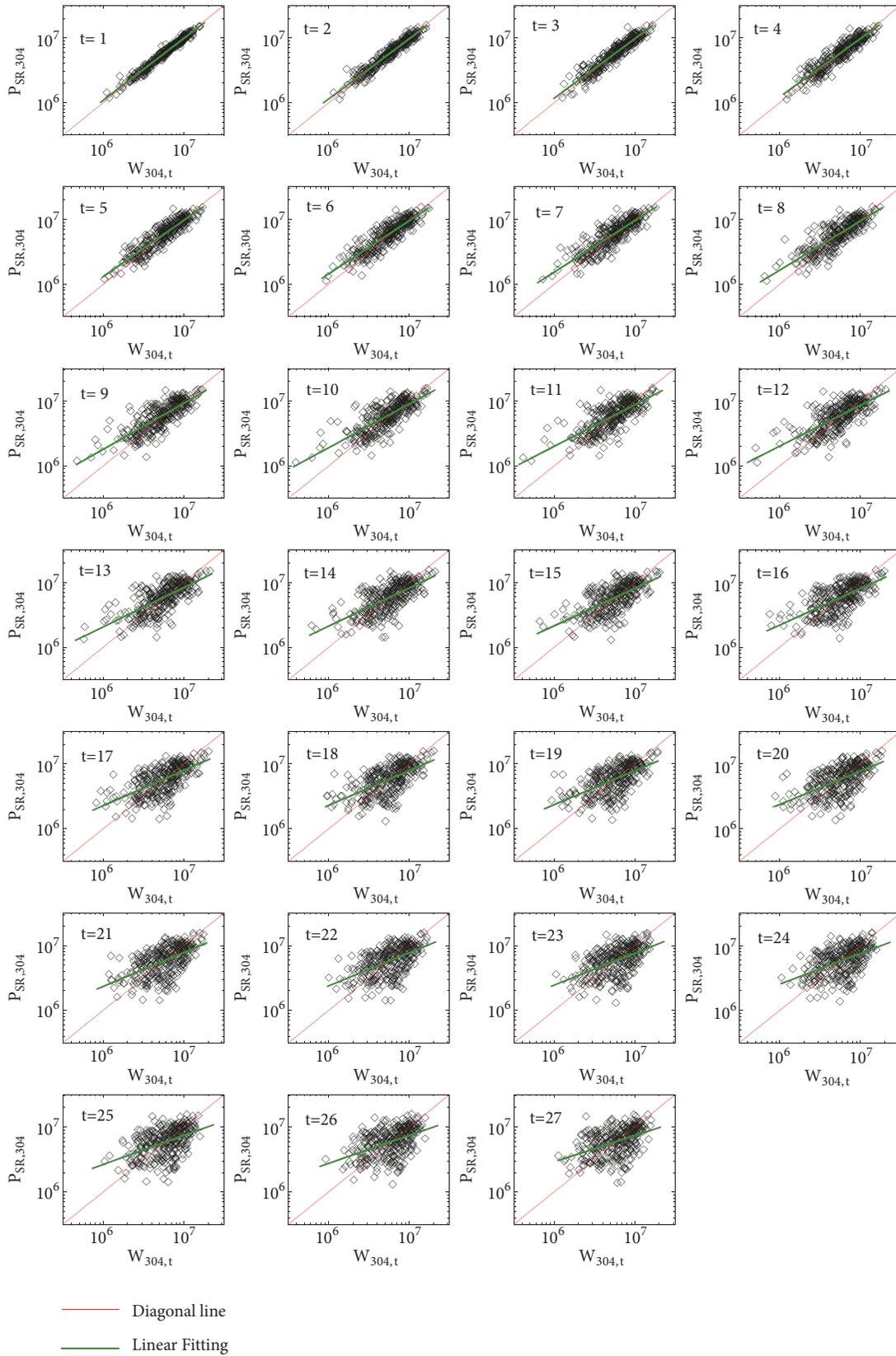


FIGURE 5: Scatter diagrams between  $W_{304,t}$  and  $P_{SR,304}$  before 27 March 2013. The solid red line is the diagonal line, and the solid green line is the fitting line.

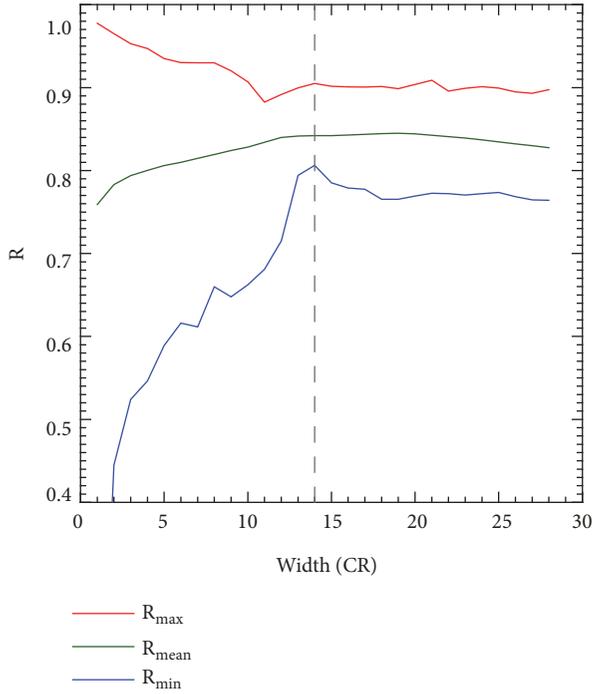


FIGURE 6: The correlation coefficient between fitted  $Y$  and  $F_{10.7}$  with different widths in moving fit from 2012 to 2013.  $R_{\min}$  (blue line),  $R_{\max}$  (red line), and  $R_{\text{mean}}$  (green line) correspond to the minimum, maximum, and average of the correlation coefficient, respectively.

The 0-day predictive and observed  $F_{10.7}$  values are compared in Figure 7. The 0-day predictive values are consistent with the observed  $F_{10.7}$  values, especially from August 2012 to August 2013. The mean relative daily forecast error is defined as  $E_{\text{mr}}^t$  in (7). The 0-day forecast relative errors ( $E_{\text{mr}}^0$ ) of our method are 5.49% in 2012, 4.86% in 2013, and 5.18% in 2012-2013.

$$E_{\text{mr}}^t = \frac{1}{N} \sum_{i=1}^N \frac{|o_{t,i} - f_{t,i}|}{o_{t,i}} \times 100\%, \quad t = 0, 1, 2, 3, \dots, 27 \quad (7)$$

where the parameter  $N$ , which represents the number of testing samples, is equal to 611 in 2012-2013. The parameter  $t$  represents the day, and  $o_{t,i}$  and  $f_{t,i}$  denote the observed and predictive values, respectively.

To show the precise of our prediction, Figure 8 shows a comparison of the 54th-order AR-model and our method with respect to the average daily predictive relative errors in advance of 1-27 days. The predictive  $F_{10.7}$  values of 54th-order AR-model are calculated with the method of Liu et al. [18].

The two kinds of  $E_{\text{mr}}^t$  increase with the growth of advance time. While the error of our model increases with a steady rate in 1-27 days, the other error grows faster in 1-9 days and approximates to horizontal lines in 10-27 days. From the contrast, it is discovered that there are two kinds of fitting errors in our model: one is in (4), and the other is in (6).  $E_{\text{mr}}^0$  in our method is the error of fitting in (4), which is up to 5.18%, and the  $E_{\text{mr}}^1$  of AR-model is below 4%. Thus, the  $F_{10.7}$  short-term predictive precision of our model is lower than that of AR-model. The reason is that the known preceding

value has the tallest weight in the predicted value of AR-model, and the  $F_{10.7}$  varies slowly and autocorrelates very strongly. Therefore, the  $E_{\text{mr}}^t$  in our method is greater than that in the AR-model for 1-3 days. The growth rate of  $E_{\text{mr}}^t$  in our method is the error of fitting in (6). Even so, because of the addition of solar back disk data in our model, the accuracy of prediction is improved obviously for 4-27 days.

There are two classical evaluation indicators used to reflect the errors of the predictive results: the mean absolute error and the mean relative error. To assess the accuracy of every testing case, the errors are defined as follows:

$$E_{\text{mr}} = \frac{1}{27} \sum_{t=1}^{27} \frac{|o_t - f_t|}{o_t} \times 100\%, \quad t = 1, 2, 3, \dots, 27 \quad (8)$$

$$E_{\text{ma}} = \frac{1}{27} \sum_{t=1}^{27} |o_t - f_t|, \quad t = 1, 2, 3, \dots, 27 \quad (9)$$

where  $o_t$  and  $f_t$  denote the observed and predictive values, respectively, and the parameter  $t$  is the day.

To assess the accuracy of all testing cases, another two kinds of error are defined in (10) and (11).

$$ME_{\text{mr}} = \frac{1}{27N} \sum_{t=1}^{27} \sum_{i=1}^N \frac{|o_{t,i} - f_{t,i}|}{o_{t,i}} \times 100\%, \quad t = 1, 2, 3, \dots, 27 \quad (10)$$

$$ME_{\text{ma}} = \frac{1}{27N} \sum_{t=1}^{27} \sum_{i=1}^N |o_{t,i} - f_{t,i}|, \quad t = 1, 2, 3, \dots, 27 \quad (11)$$

where the parameter  $N$  represents the number of testing samples,  $t$  represents the day, and  $o_{t,i}$  and  $f_{t,i}$  denote the observed and predictive values, respectively.

To assess the drop-rate between our model and AR-model, the parameters  $DR_{\text{mr}}$  and  $DR_{\text{ma}}$  are defined as follows:

$$DR_{\text{mr}} = \frac{ME'_{\text{mr}} - ME''_{\text{mr}}}{ME'_{\text{mr}}} \times 100\% \quad (12)$$

$$DR_{\text{ma}} = \frac{ME'_{\text{ma}} - ME''_{\text{ma}}}{ME'_{\text{ma}}} \times 100\% \quad (13)$$

where the  $ME'_{\text{mr}}$  and  $ME'_{\text{ma}}$  represent  $ME_{\text{mr}}$  and  $ME_{\text{ma}}$  of AR-model, respectively, during 2012-2013.  $ME''_{\text{mr}}$  and  $ME''_{\text{ma}}$  represent the  $ME_{\text{mr}}$  and  $ME_{\text{ma}}$  of our model, respectively, during 2012-2013.

Figure 9 shows the observed values of  $F_{10.7}$  and the values of  $E_{\text{mr}}$  in the 54th-order AR-model and our method during 2012-2013. Due to the uncontrolled spin of satellites, instrument failure, or their inappropriate relative positions, SDO and the twin STEREO spacecraft cannot capture the full-disk EUV images clearly and completely, which should be omitted from the testing sample set (yellow vertical short line at the top in Figure 9). Finally, there are 611 samples for testing during 2012-2013, with 304 samples in 2012 and 307 samples in 2013.

The tendencies of  $E_{\text{mr}}$  in the two methods are approximated in Figure 9. The  $E_{\text{mr}}$  of our method is slightly less than

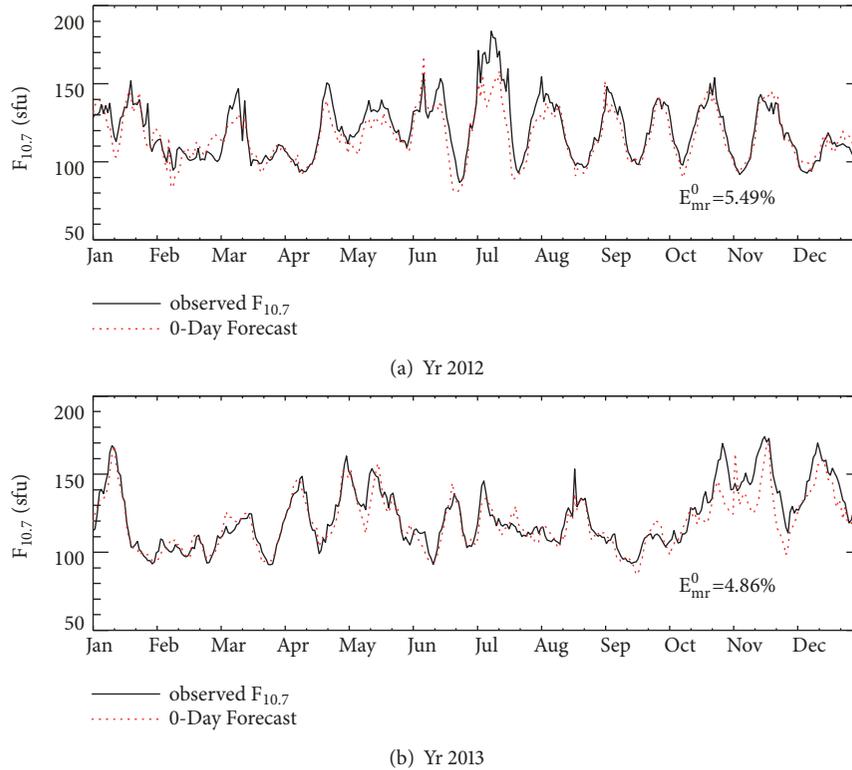


FIGURE 7: A comparison of 0-day predictive (dotted red line) and observed (solid black line)  $F_{10.7}$  in 2012 (a) and 2013 (b).

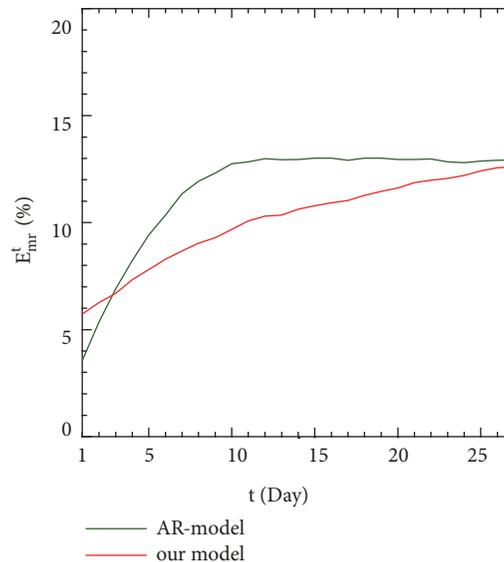


FIGURE 8: The  $E_{mr}^t$  of AR-model (green line) and our method (red line) during 2012-2013 for 1-27 days.

that of AR-model, especially during June 2012 and October 2013. A threshold of 10%  $E_{mr}$  is defined here to evaluate the accuracy of the forecasting results.

The  $E_{mr}$  in our method is less than 10% from July to November in 2012, which means that the prediction of our method is satisfied when the period of  $F_{10.7}$  is stable. Compared with the  $ME_{mr}$  in the 54th-order AR-model, that

in our method decreases from 11.54% to 10.09% with a 12.52% drop-rate during 2012-2013 (shown in Table 2). The drop-rate of  $ME_{ma}$  is up to 10.35%. Table 2 shows that the accuracy of our method is desirable during testing. Additionally, two forecasting cases are chosen for further analysis in Figure 10.

The  $E_{mr}$  in our method is much less than that of the 54th-order AR-model in the case of 25 March 2013. Figure 10(a)

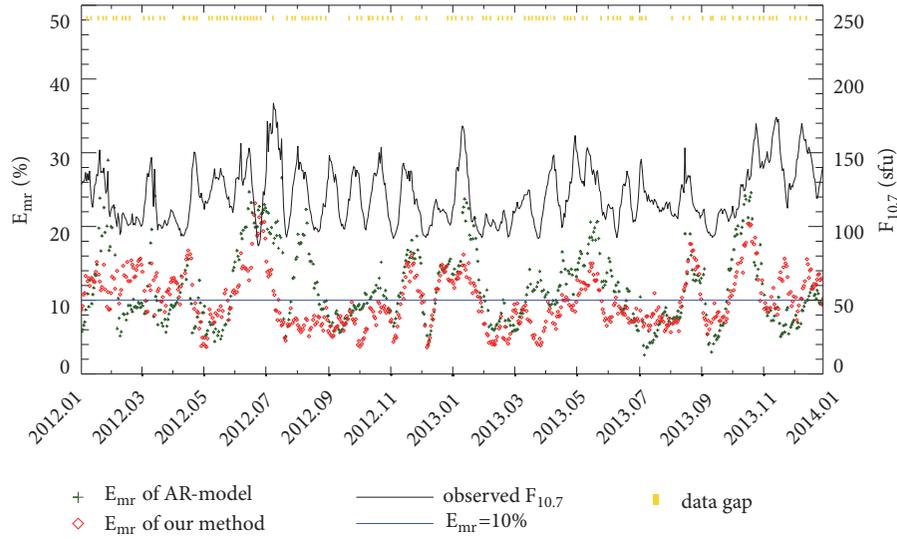


FIGURE 9:  $E_{mr}$  derived by the 54th-order AR-model (green plus signs) and our method (red diamonds) during 2012-2013. The solid black line represents observed  $F_{10.7}$ , and the solid blue line represents  $E_{mr}$  equal to 10%. The short yellow vertical line at the top represents missing data due to the incomplete EUV images.

TABLE 2: The  $ME_{ma}$  and  $ME_{mr}$  values of AR-model and our method during 2012-2013.

		Year			Drop-rate <sup>a</sup> (%)
		2012	2013	2012-2013	
$ME_{mr}$ (%)	AR-model	11.91	11.17	11.54	
	Our method	10.28	9.91	10.09	12.52
$ME_{ma}$ (sfu)	AR-model	14.25	14.09	14.17	
	Our method	12.67	12.74	12.71	10.35

<sup>a</sup> The drop-rate is defined in (12) and (13).

shows a scatter diagram of observed  $P_{SR,304}$  and  $F_{10.7}$  in the previous 14 CRs and the future 27 days. The fitting line in Figure 10(a) is fitted by the data of the previous 14 CRs. The correlation coefficient (R) between fitted Y and  $F_{10.7}$  is as high as 0.92. Figure 10(b) shows a comparison of observed  $F_{10.7}$ ,  $P_{SR,304}$ ,  $W_{304,t}$  and  $S_{304,t}$  in the previous and future 27 days from the test date, 25 March 2013. Figure 10(c) shows a comparison of observed and predictive  $F_{10.7}$ . The results of our method are closer to the observed  $F_{10.7}$  than that of the 54th-order AR-model. The ARs (NOAA 11711, 11715, 11716, 11717, and 11718) on the Earth-side SDO/AIA EUV images on 8 April 2013 (Figure 11(c)) were already captured 14 days in advance by the EUV images on 25 March 2013 (Figure 11(a)). Additionally, these ARs move towards the Earth without disappearance, and no new ARs appear. In summary, when the ARs on the far-side EUV images change slowly in the next Carrington rotation, the forecast accuracy of our method is preferable to that of AR-model. This is the greatest advantage of our model in comparison to the AR-model.

Another case on 28 February 2012 is shown in Figures 10(d), 10(e), and 10(f). There is a crest of  $F_{10.7}$  from 1 to 14 March 2012 along with a trough of  $W_{304,t}$  and  $S_{304,t}$  (see

Figure 10(e)). Contrasting the three images in Figure 12, new ARs (NOAA 11429 and 11430) appear, but they are too weak in the solar full-disk EUV image (Figure 12(a)). These ARs then strengthen while they rotate towards the Earth on 8 March 2012 (Figure 12(c)).  $E_{mr}$  of our method decreases from 15.65% on 28 February 2012 to 8.45% on 6 March 2012, as seen in Figure 13. Our model could recognize the approximate position of NOAA 11429 and 11430 on 4 March 2012 because the forecast values of  $F_{10.7}$  remain close to the observed values after 14 March 2012 (corresponding to 10, 9, and 8 on horizontal axis in Figures 13(d), 13(e), and 13(f), respectively). Thus, once the solar full-disk EUV image captures the new SRs, the forecast results of our model are adjusted rapidly, and the predictive precision is improved quickly. This is the second advantage of our model in comparison to the AR-model.

In comparison with AR-model in Figures 10(c) and 10(f), the accuracy of prediction is unsatisfactory in the coming 1-2 days. This finding agrees with the results in Figure 8. Thus, to gain a more satisfactory predictive precision, we can combine the upcoming 1-2 days predictive values of the 54th-order AR-model with the upcoming 3-27 days predictive values of our method in practice.

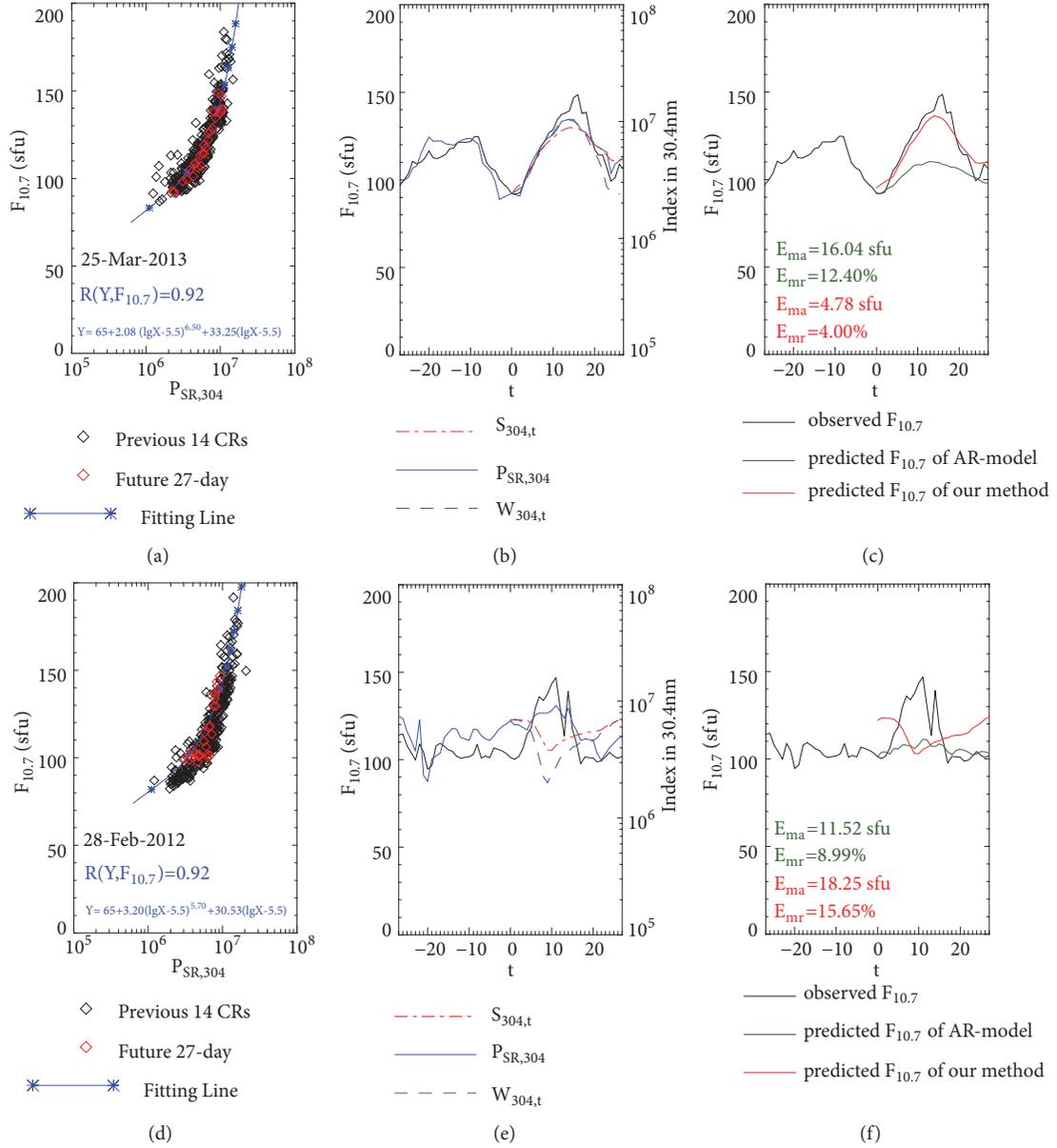


FIGURE 10: The  $F_{10.7}$  forecast results of two cases: 25 March 2013 (a, b, c) and 28 February 2012 (d, e, f). (a, d) Scatter diagrams of the 14 CRs' previous values of  $F_{10.7}$  and  $P_{SR,304}$  (black diamonds), and the future 27-day values of  $F_{10.7}$  and  $P_{SR,304}$  (red diamonds). The solid blue line with asterisks is the fitted function.  $R$  is the correlation coefficient between fitted  $Y$  and  $F_{10.7}$ . (b, e) Comparison of  $P_{SR,304}$  (solid blue line),  $W_{304}$  (dashed blue line),  $S_{304}$  (dashed-dot red line), and  $F_{10.7}$  (solid black line) in previous (negative value in horizontal axis) and future (positive value in horizontal axis) 27 days of that test date (zero in horizontal axis). (c, f) The value of observed  $F_{10.7}$  (black line) and predictive  $F_{10.7}$  obtained from the 54th-order AR-model (green line) and our method (red line) in future (positive value in the horizontal axis) 27 days.

#### 4. Conclusion

Through the error analysis and case studies on the medium-term forecast experiments of daily  $F_{10.7}$  during 2012-2013, we can obtain the following conclusions.

The full-disk EUV images can provide coronal information about the far-side solar disk, which is 13.5 days earlier than that of other models using only Earth-side information or  $F_{10.7}$  itself. Thus, the  $F_{10.7}$  medium-term forecast accuracy

of our method is better than that of the 54th-order AR-model, especially for upcoming 3-27 days.

The sensitivity of our model is much higher than that of AR-model. The forecast results of our model can adjust rapidly, and the predictive precision is improved quickly. The drop-rate of  $ME_{mr}$  in our method is 12.52% during 2012-2013.

This was the first attempt in which we found a proxy in solar EUV images to represent the coronal contribution to  $F_{10.7}$ . It was also the first attempt to forecast the upcoming

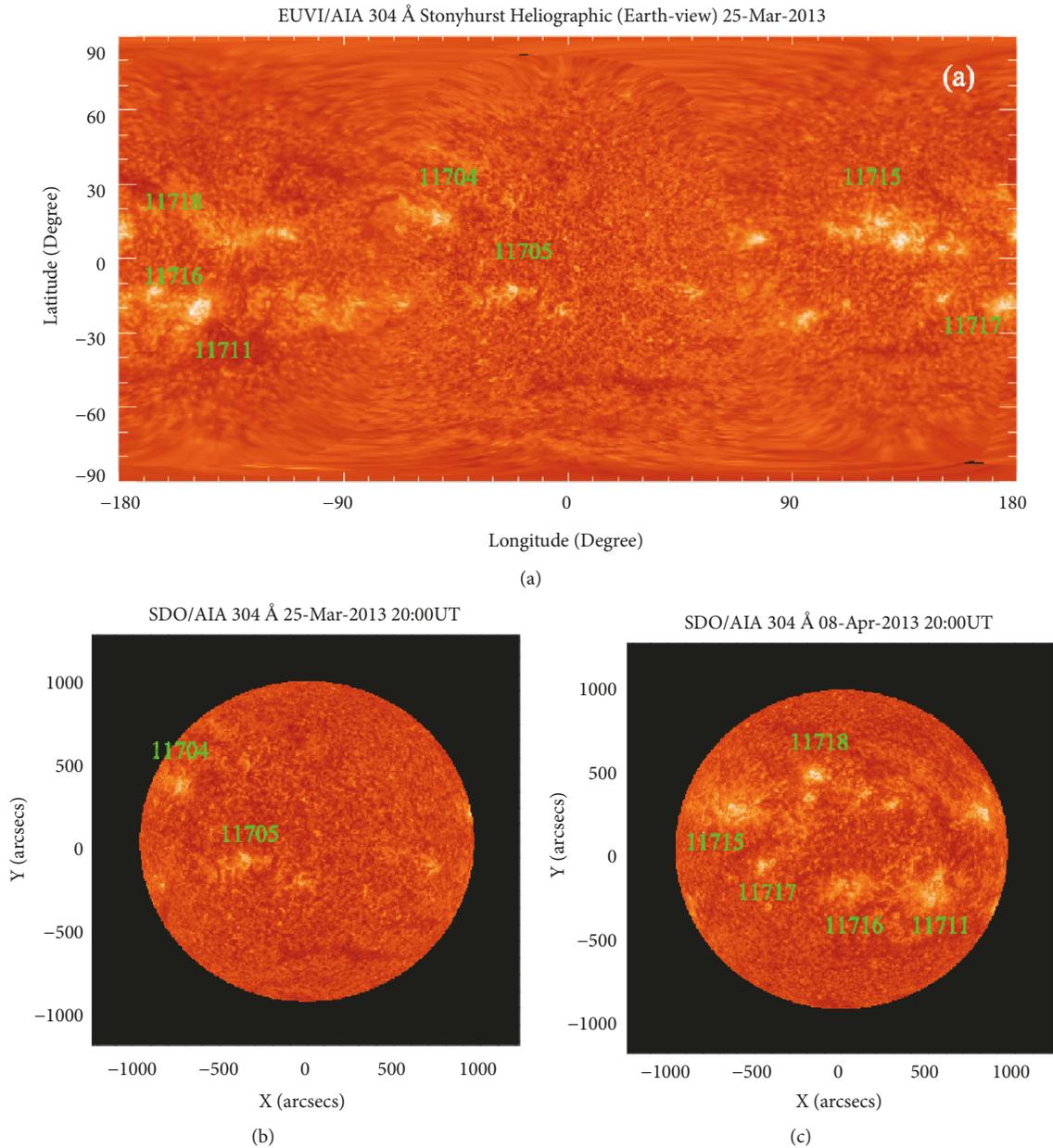


FIGURE 11: The 304 Å full-disk EUV image on 25 March 2013 (a) and Earth-side EUV images on 25 March 2013 (b) and 8 April 2013 (c).

27-day values of  $F_{10.7}$  based on the solar full-disk EUV images. Moreover, to gain a more satisfactory predictive precision, combining the upcoming 1-2 days predictive values of the 54th-order AR-model with the upcoming 3-27 days predictive values of our method in practice should be considered.

Although there are problems with receiving real-time data from satellites STEREO/EUVI and SDO/AIA, this paper importantly demonstrates the tangible benefits that 360 degree solar observations provide for the prediction of solar activity. The Lagrangian 5 (L5) point lies at the third corners of the equilateral triangles in the plane of orbit whose common base is the line between the centers of Sun and

Earth. So the viewing angle in L5 point can reach -150 degrees in Stonyhurst heliographic coordinates and the L5 observing platform can provide the EUV images before about 11.25 day. If the current data could be afforded by the ability of the L5 observing platform, it would enable this technique to forecast about 11.25 days  $F_{10.7}$  in practice.

### Data Availability

All data used in the manuscript can be downloaded from the available database of the websites. And below are these URL of websites. (1) The  $F_{10.7}$  index data used to support the findings of this study can be downloaded from the

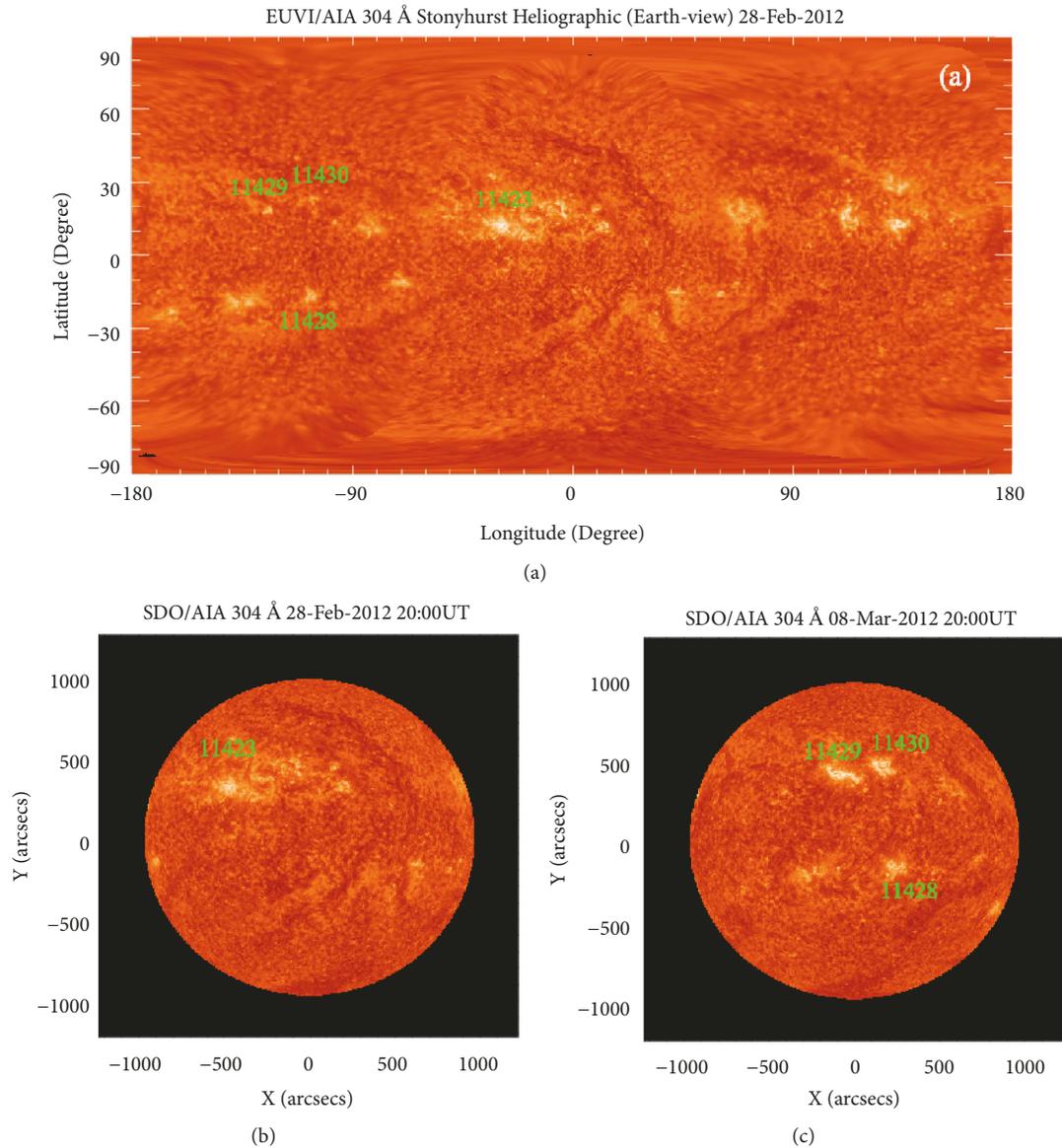


FIGURE 12: The 304 Å full-disk EUV images on 28 February 2012 (a) and Earth-side EUV images on 28 February 2012 (b) and 8 March 2012 (c).

available database of the National Oceanic and Atmospheric Administration (NOAA) ([ftp://ftp.ngdc.noaa.gov/STP/space-weather/solar-data/solar-features/solar-radio/noontime-flux/penticton/penticton\\_observed/listings/listing\\_drao\\_noontime-flux-observed\\_daily.txt](ftp://ftp.ngdc.noaa.gov/STP/space-weather/solar-data/solar-features/solar-radio/noontime-flux/penticton/penticton_observed/listings/listing_drao_noontime-flux-observed_daily.txt)). (2) The daily level-1 FITS (Flexible Image Transport System) files of SDO/AIA from May 2010 to December 2015 are downloaded from the available database of the Joint Science Operations Center (JSOC) at Stanford University (<http://jsoc.stanford.edu/>). (3) The daily FITS files of STEREO/EUVI from January 2011 to December 2013 are downloaded from the available database of STEREO Science Center (<https://stereoftp.nascom.nasa.gov/data/beacon/ahead/secchi/img/euvi/> and <https://stereoftp.nascom.nasa.gov/data/beacon/behind/secchi/img/euvi/>).

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The 10.7cm solar radio flux data are provided as a service by the National Research Council of Canada. We thank the SDO/AIA instrument team for providing the coronal observations and Stanford University's JSOC for providing the available database. We also thank the STEREO/EUVI instrument team and NOAA of the USA for providing their data. This work was supported by the National Defense Science and Technology Innovation Special Zone, the National

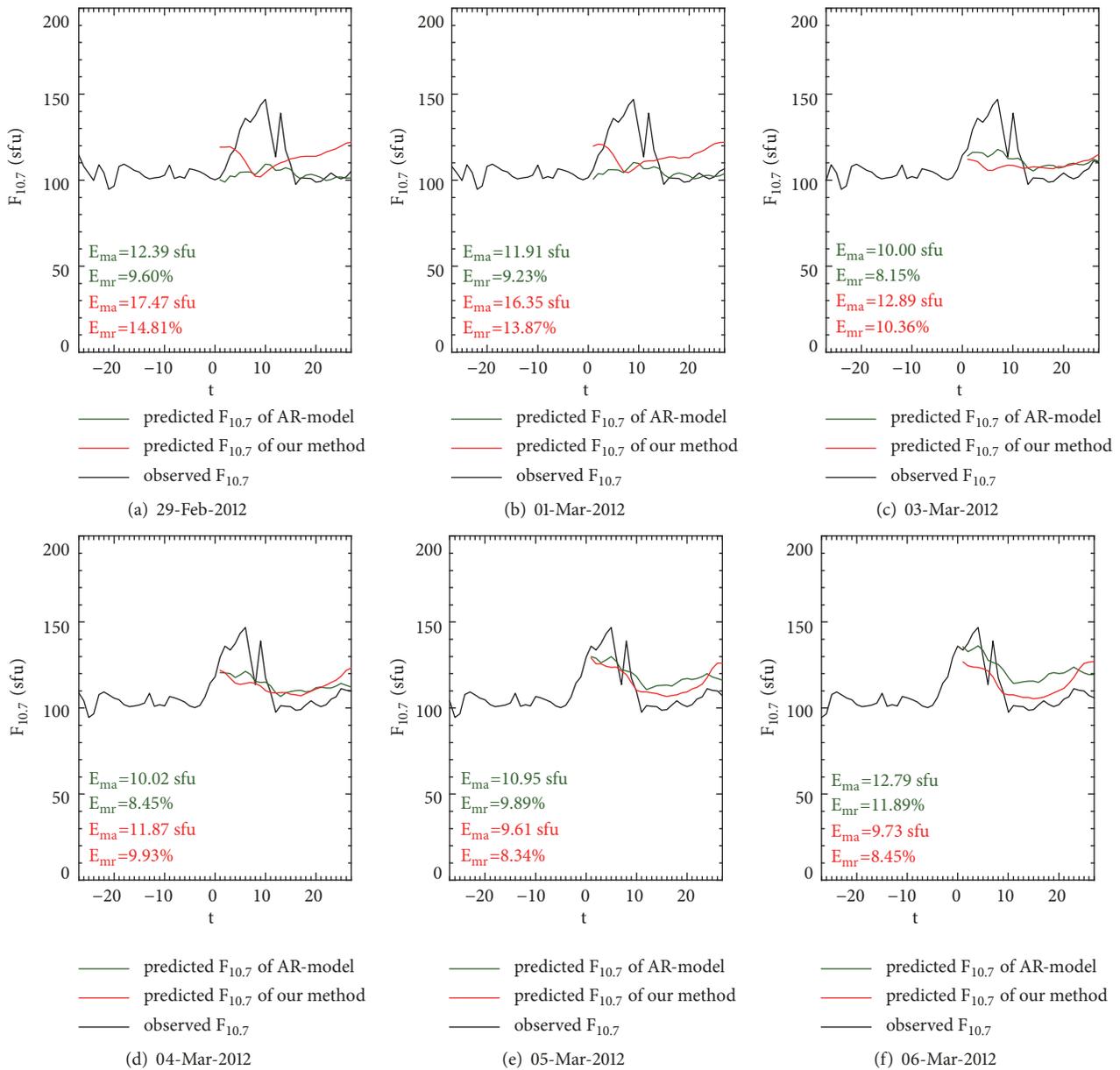


FIGURE 13: A comparison of observed  $F_{10.7}$  (black line) and predictive  $F_{10.7}$  obtained from AR-model (green line) and our method (red line) from 29 February 2012 to 6 March 2012.

Natural Science Foundation of China (Grant no. Y75037A070 and Grant no. 41604149), the Beijing Municipal Science and Technology Project (project number Z181100002918004), and the Strategic Priority Research Program of Chinese Academy of Sciences (Grant no. XDA17010302).

## References

- [1] K. O. Kiepenheuer, "Structure and development of solar active regions," *International Astronomical Union*, vol. 35, no. 379, 1968.
- [2] H. Schunker and P. S. Cally, "Magnetic field inclination and atmospheric oscillations above solar active regions," *Monthly Notices of the Royal Astronomical Society*, vol. 372, no. 2, pp. 551–564, 2006.
- [3] R. G. Athay and J. M. Beckers, "The solar chromosphere and corona: quiet sun," *Physics Today*, vol. 29, no. 11, pp. 74–76, 1976.
- [4] J. Lee, "Radio emissions from solar active regions," *Space Science Reviews*, vol. 133, no. 1-4, pp. 73–102, 2007.
- [5] S. J. Schonfeld, S. M. White, C. J. Henney, C. N. Arge, and R. T. J. McAteer, "Coronal sources of the solar  $F_{10.7}$  radio flux," *The Astrophysical Journal*, vol. 808, no. 1, 2015.
- [6] W. N. Christiansen, D. E. Yabsley, and B. Y. Mills, "Measurements of solar radiation at a wavelength of 50 centimetres during the eclipse of november 1, 1948," *Australian Journal of Chemistry*, vol. 2, no. 4, pp. 506–523, 1949.

- [7] W. Livingston, M. J. Penn, and L. Svalgaard, "Decreasing sunspot magnetic fields explain unique 10.7cm radio flux," *The Astrophysical Journal Letters*, vol. 757, no. 1, 2012.
- [8] C. L. Selhorst, J. E. Costa, C. G. Giménez de Castro et al., "The 17 GHz active region number," *The Astrophysical Journal*, vol. 790, no. 2, article 134, 2014.
- [9] K. F. Tapping, "The 10.7 cm solar radio flux ( $F_{10.7}$ )," *Space Weather-the International Journal of Research & Applications*, vol. 11, no. 7, pp. 394–406, 2013.
- [10] W. K. Tobiska, "Validating the solar EUV proxy, E10.7," *Journal of Geophysical Research: Space Physics*, vol. 106, no. A12, pp. 29969–29978, 2001.
- [11] A. E. Hedin, "The atmospheric model in the region 90 to 2000 km," *Advances in Space Research*, vol. 8, no. 5-6, pp. 9–25, 1988.
- [12] J. M. Picone, A. E. Hedin, D. P. Drob, and A. C. Aikin, "NRLMSISE-00 empirical model of the atmosphere: statistical comparisons and scientific issues," *Journal of Geophysical Research: Space Physics*, vol. 107, no. A12, pp. SIA15-1–SIA15-16, 2002.
- [13] S. H. Knowles, J. M. Picone, S. E. Thonnard, and A. C. Nicholas, "The effect of atmospheric drag on satellite orbits during the bastille day event," *Solar Physics*, vol. 204, no. 1-2, pp. 387–397, 2001.
- [14] D. H. Hathaway, "The solar cycle," *Living Reviews in Solar Physics*, vol. 7, no. 1, 2010.
- [15] R. Vautard, P. Yiou, and M. Ghil, "Singular-spectrum analysis: a toolkit for short, noisy chaotic signals," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1–4, pp. 95–126, 1992.
- [16] H. Akaike, "Fitting autoregressive models for prediction," *Annals of the Institute of Statistical Mathematics*, vol. 21, pp. 243–247, 1969.
- [17] Z. Q. Zhong, Q. S. Liu, X. J. He, and J. C. Gong, "Application of singular spectrum analysis to solar 10.7cm radio flux 27-day forecast," *Chinese Journal of Space Science*, vol. 25, no. 3, pp. 199–203, 2005.
- [18] S. Q. Liu, Q. Z. Zhong, J. Wen, and X. K. Dou, "Modeling research of the 27-day forecast of 10.7 cm solar radio flux (I)," *Chinese Astronomy and Astrophysics*, vol. 34, no. 3, pp. 305–315, 2010.
- [19] H. B. Wang, J. N. Xiong, and C. Y. Zhao, "The mid-term forecast method of solar radiation index," *Chinese Astronomy and Astrophysics*, vol. 39, no. 2, pp. 198–211, 2015.
- [20] J. Wen, Z. Q. Zhong, and S. Q. Liu, "Model research of 10.7 cm solar radio flux 27-day forecast," *Chinese Journal of Space Science*, vol. 30, no. 3, article 198, 2010.
- [21] C. J. Henney, W. A. Toussaint, S. M. White, and C. N. Arge, "Forecasting  $F_{10.7}$  with solar magnetic flux transport modeling," *Space Weather Journal*, vol. 10, article S02011, 2012.
- [22] E. Quémerais and J. L. Bertaux, "14-day forecast of solar indices using interplanetary Lyman  $\alpha$  background data," *Geophysical Research Letters*, vol. 29, no. 2, pp. 5-1–5-4, 2002.
- [23] J. L. Bertaux, E. Kyrölä, E. Quémerais et al., *SWAN: A Study of Solar Wind Anisotropies on SOHO with Lyman Alpha Sky Mapping*, Springer, Netherlands, 1995.
- [24] C. Lindsey and D. Braun, "Seismic imaging of the sun's far hemisphere and its applications in space weather forecasting," *Space Weather-the International Journal of Research & Applications*, vol. 15, 2017.
- [25] B. O'Dwyer, G. Del Zanna, H. E. Mason, M. A. Weber, and D. Tripathi, "SDO/AIA response to coronal hole, quiet Sun, active region, and flare plasma," *Astronomy & Astrophysics*, vol. 521, article A21, 2010.
- [26] K. F. Tapping, D. Boteler, P. Charbonneau et al., "Solar magnetic activity and total irradiance since the maunder minimum," *Solar Physics*, vol. 246, no. 2, pp. 309–326, 2007.
- [27] J. R. Lemen, A. M. Title, D. J. Akin et al., "The atmospheric imaging assembly (aia) on the solar dynamics observatory (sdo)," *Solar Physics*, vol. 275, no. 1-2, pp. 17–40, 2012.
- [28] N. Balan, G. Bailey, and Y. Su, "Variations of the ionosphere and related solar fluxes during solar cycles 21 and 22," *Advances in Space Research*, vol. 18, no. 3, pp. 11–14, 1996.
- [29] K. Tapping and C. Morgan, "Changing relationships between sunspot number, total sunspot area and F10.7 in cycles 23 and 24," *Solar Physics*, vol. 292, no. 6, 2017.
- [30] R. A. Howard, J. D. Moses, D. G. Socker, K. P. Dere, and J. W. Cook, "Sun earth connection coronal and heliospheric investigation (SECCHI)," *Space Science Reviews*, vol. 136, no. 1-4, pp. 67–115, 2008.
- [31] C. Lowder, J. Qiu, R. Leamon, and Y. Liu, "Measurements of EUV coronal holes and open magnetic flux," *The Astrophysical Journal*, vol. 783, no. 142, p. 13, 2014.
- [32] J. E. Vernazza, E. H. Avrett, and R. Loeser, "Structure of the solar chromosphere. III - Models of the EUV brightness components of the quiet-sun," *The Astrophysical Journal Supplement Series*, vol. 45, no. 4, pp. 635–725, 1981.
- [33] L. D. Krista and P. T. Gallagher, "Automated coronal hole detection using local intensity thresholding techniques," *Solar Physics*, vol. 256, no. 1-2, pp. 87–100, 2009.
- [34] D. Pérez-Suárez, P. Higgins A, D. Bloomfield S et al., "Automated solar feature detection for space weather applications," *Applied Signal and Image Processing: Multidisciplinary Advancements*, pp. 207–225, 2011.
- [35] J. H. Zar, "Significance testing of the spearman rank correlation coefficient," *Journal of the American Statistical Association*, vol. 67, no. 339, pp. 578–580, 1972.
- [36] S. B. Lyerly, "The average Spearman rank correlation coefficient," *Psychometrika*, vol. 17, no. 4, pp. 421–428, 1952.
- [37] P. C. Liewer, J. Qiu, and C. Lindsey, "Comparison of helioseismic far-side active region detections with STEREO far-side euV observations of solar activity," *Solar Physics*, vol. 292, no. 10, article 146, 2017.