The Astrophysical Journal, 902:72 (9pp), 2020 October 10 https://doi.org/10.3847/1538-4357/abb2a0

A New Approach for the Regression of the Center Coordinates and Radius of the Solar Disk Using a Deep Convolutional Neural Network

Gaofei Zhu[1,2], Ganghua Lin[1], Dongguang Wang[1], and Xiao Yang[1]

[1] Key Laboratory of Solar Activity, National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100012, People's Republic of China

[2] University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China

Received 2020 July 12; revised 2020 August 22; accepted 2020 August 24; published 2020 October 13

Abstract

This paper presents a new approach for the regression of the center coordinates and radius of the solar disk in Hα solar full-disk images by using a Deep Convolutional Neural Network. We use ~100,000 original Hα solar full-disk images obtained from Huairou Solar Observing Station as the experimental data set. The data set includes two parts: the original image and three numeric values (center coordinates and radius). In order to deal with the uneven distribution of the solar disk position in the original image, we randomly shift the solar disk during image preprocessing. Furthermore, data augmentation is also used to increase the robustness of the model. By evaluating the model with R-square and relative error, the center coordinates and the radius of the solar disk are proved to be effectively regressed. The data sets we constructed and source code are available as open source on GitHub.

Unified Astronomy Thesaurus concepts: Solar radius (1488); Solar physics (1476); Astronomy data analysis (1858); Convolutional neural networks (1938); Regression (1914)

1. Introduction

Hα (hydrogen-alpha) solar full-disk observation, which presents solar chromospheric activities, is an essential record in solar physics. Sunspots, flares, filaments (prominences), spicules, and fibrils are the phenomena easily observed in the chromosphere (Figure 1). There are many observatories around the world performing patrol Hα observations of the Sun. For example, the Global Oscillation Network Group (Harvey et al. 1996) consists of six observing stations (Hill et al. 1994): Big Bear Solar Observatory, Mauna Loa Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Teide Observatory, and Gerro Tololo Intercontinental Solar Observatory. Besides, some other observatories are observing daily the solar chromosphere, such as Huairou Solar Observing Station in China, which is one of the global high-resolution Halpha network.

Large amounts of observation data have been accumulated during long-term observations. There exist various recognition studies on solar activity phenomena from Hα solar full-disk images such as filaments (Yuan et al. 2011; Hao et al. 2015; Zhu et al. 2019; Lin et al. 2020), flares (Pötzi et al. 2015; Yang et al. 2018), and sunspots (Zharkov et al. 2005; Curto et al. 2008; Watson et al. 2009; Goel & Mathew 2014; Zhao et al. 2016). A crucial step in the study of solar activity in the raw Hα solar full-disk images is to obtain the center coordinates and radius of the solar disk, which is an important basis for subsequent processing. The distance between the Earth and the Sun, as well as the stability of the telescope, result in different center coordinates and radius values. With the introduction of image processing technology, the most commonly used methods are currently divided into two branches. One is using the least-squares to perform the circle fitting through the inflection points of the edge of the solar disk (Veronig et al. 2000). The other is adopting the Hough transform (Ballard 1981; Gonzalez & Woods 2007), which is often used in image processing for circle detection, however, it is inefficient for identifying solar disks. In order to reduce the memory consumption and to speed up the running time, Yuan et al. (2011) use the so-called cascading Hough circle detector which resizes the original image by the factor of $1/k$, then does the Hough transform, and at last restores the center and radius of the original solar disk accordingly. Hao et al. (2015) combined the two methods in the study of solar filament detection in order to accurately determine and check the center location and the radius of the solar disk. If the difference between the results of the two methods is greater than a certain threshold, the results will be considered failed, and manual inspection is required. Although both methods can obtain the center coordinates and radius of the solar disk, they still have certain limitations. After filtering and edge detection, for the least-squares circle fitting, we need to exclude noise interference from the inside and outside of the solar disk as much as possible to obtain the pixels of a clean solar edge. Only in this way the accurate center and radius can be obtained. However, in actual operation, it is often difficult to rule out all interference factors. Especially not all inflection points are on the circle because of the lower corona that is visible in Hα. For the Hough transform circle detection method, which is not sensitive to noise, the requirements of its initial parameters are relatively high, the radius range and step size need to be set. Furthermore, the amount of calculations is huge for the large-size solar full-disk images. Although the calculation amount can be reduced to a certain extent by using a cascading Hough circle detector (Yuan et al. 2011), the accuracy of the obtained center and radius of the solar disk will decrease to some extent as the image is reduced.

With the development of computer hardware and more and more data being generated, the deep Convolutional Neural Network (CNN) has begun to show strong vitality in various scientific fields including astrophysical problems. Cabrera-Vives et al. (2017) presented a rotation-invariant CNN model, named Deep-HiTS, to classify the images from the High cadence Transient Survey into real and fake transients. Zheng et al. (2019) used two CNN models to predict the occurrence of four classes of solar flares (i.e., No-flare, C, M, and X) within

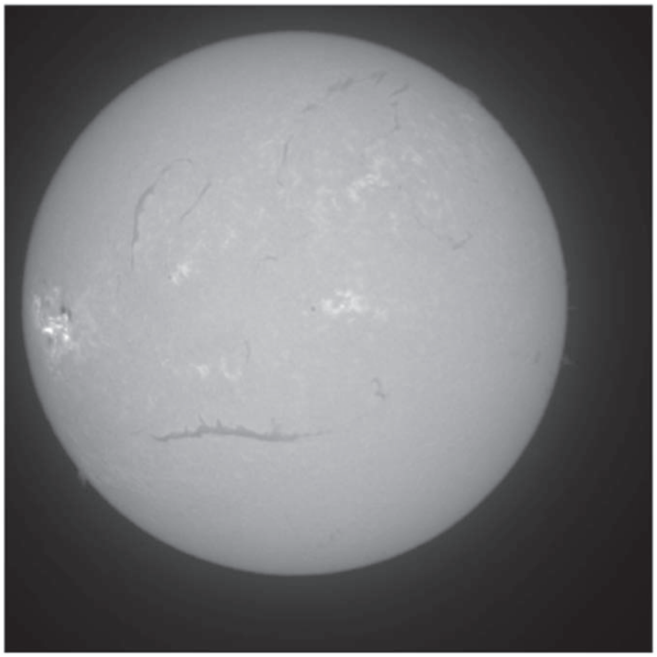

Figure 1. Hα full-disk solar image. There are several solar activities can be observed in the solar chromosphere, such as flares, sunspots, fibrils, and filaments.

24 hr. Park et al. (2018) and Li et al. (2020) also adopt a variety of CNN models to predict solar flare occurrence. Lanusse et al. (2018) trained a CNN to identify the existence of gravitational lensing in simulated data. Van Oort et al. (2019) presented a convolutional approach called CASI for shell identification, and the data they used were also simulated. Among so many astrophysical studies using deep CNNs, most of them are mainly focused on the classification, recognition and detection of specific targets. In addition, there is a part of the research focused on regression applications using CNN models. Ho et al. (2019) presented two CNN models for dynamical mass measurements of Galaxy Clusters. However, there is an important issue that cannot be ignored: most of the data sets used in current researches are simulated data, not true observational ones. As Van Oort et al. (2019) mentioned in their conclusions, there is still a huge difference between the models they trained with simulated data and true observational data.

In this paper, we design a CNN network, which can learn relevant features on target objects from raw true observation data and regress automatically the particular values of the center coordinates and radius of the solar disk from the learned features in Hα solar full-disk images. This paper is organized as follows, Section 2 is focused on data, the proposed model is given in Section 3, results are presented in Section 4, and finally, discussions and conclusions are provided in Section 5.

2. Data

2.1. Data Acquisition

Since there is currently no data set available for our research, we obtain nearly 100,000 Hα solar full-disk images from the Huairou Solar Observing Station (HSOS), National Astronomical Observatories, Chinese Academy of Sciences. HSOS is one of the world's main stations dedicated to observing the Sun. It observes the photosphere and chromosphere of the Sun at the same time and provides a variety of regional and full-disk routine observation data, such as magnetograms, filtergrams, and dopplergrams. Before 2006, a 14cm-aperture optical telescope that was one channel telescope of the Solar Multi-Channel Telescope (Ai & Hu 1986; Deng et al. 1997) was used for observing the full-disk chromosphere with a tunable Hα filter (−32 Å–+32 Å with bandwidth 0.5 Å) and a KODAK MegaPlus 4.2i Camera. After 2006, a new 20 cm-aperture telescope, which is one channel of the Solar Magnetism and Activity Telescope (Zhang et al. 2007), was built, continuing the Hα full-disk observation. Its wavelength is 6562.81 Å, with the bandpass of 0.25 Å. The center wavelength of the filter can be tuned within ±2 Å from the line center (Liu 2020). The spatial resolution, cadence, and the size of the full-disk Hα filtergrams are about 2 arcsec, 1-5 min, and 2712 × 2712, respectively.

Our data set contains two parts, one is the original images, and the other is the numerical center coordinates and radius stored in the "txt" file. We use the Hough transform and the least-squares method to automatically calculate unverified circle center coordinates and radius for producing the data set. We draw the circles on the original images, and delete the samples with deviation. The original images have to be resized to 512 × 512 due to the memory limit of the experimental 1070Ti GPU.

In daily observations, the solar disk is not always in the center of the image, and it has a tendentious position in the image during a period of time, which is not random. To increase the generalization of our model, we adopt a strategy of randomizing the center location of the solar disk. We randomly shift the solar disk with the image border cut or filled, and make sure the full disk is inside the image without being cropped.

2.2. Data Augmentation

Data augmentation is a common and effective way to improve the generalization ability and robustness of the model. It increases both the amount of training data and the noise data. There are several forms of data augmentation, such as flip, rotation, scale, crop, translation, and noise. Especially for image classification tasks, a model will be implemented using multiples of them. The atmosphere and clouds are the main interference factors for ground observations, which result in bad quality of the data, these factors are not considered in our model training. In order to demonstrate that our approach can still effectively regress the required parameters, when the image is acceptably contaminated, we add salt-and-pepper noise to the image, which generates white dots and black dots randomly in the image. The radius scaling of the solar disk is also included in our code as the visual radius variation due to the Sun–Earth distance is intrinsically present in the original data set.

After randomly generating the center of the solar disk, we add a noise density up to 0.1 to the images in each batch using a step size of 0.01. Figure 2 shows an example of different noise densities. This allows the model to obtain effective center coordinates and the radius of the solar disk in the presence of noise.

3. Proposed Model

In this section, we present a new approach using the deep CNN to obtain the center coordinate and radius of the solar disk for Hα solar full-disk image. It can regress the center coordinate

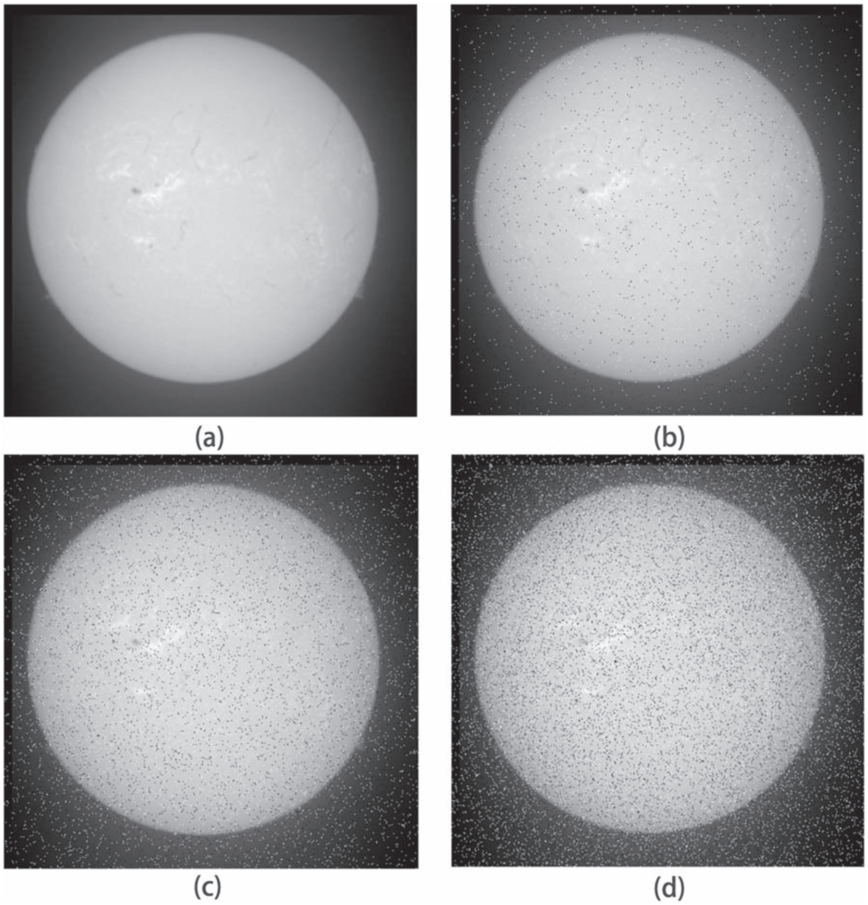

Figure 2. An example of data augmentation. (a) Only an image of randomized center coordinates without any noise. (b) An image of randomized center coordinates with a noise density of 0.01. (c) An image of randomized center coordinates with a noise density of 0.04. (d) An image of randomized center coordinates with a noise density of 0.09.

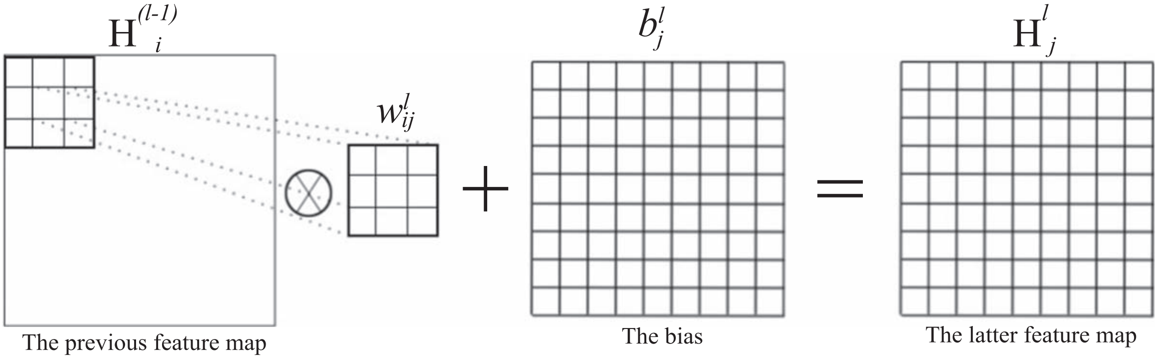

Figure 3. A basic image convolutional layer model. The feature map of H_j^l is related to each feature map of the previous layer.

and radius of the solar disk according to the learned image feature automatically. In Section 3.1, the basic conception of CNN is given. The proposed network architecture is described in Section 3.2. Finally, we describe the details of network training.

3.1. Convolutional Neural Network

The CNN first proposed by Lecun et al. (1998) has begun to show strong vitality in various scientific fields including astrophysics, and has shown its huge advantages of automatic learning capabilities. CNN can automatically learn features from a large amount of raw data without artificially adding any exact prior knowledge, while the traditional image processing requires certain prior understanding of image features. Traditional image processing often requires designing different descriptors to extract various features, such as edges, corners, and areas. Nevertheless, CNNs can learn feature classifiers automatically from the samples. It has accomplished amazing results on many issues, especially on tasks such as image classification, object detection, and segmentation.

A typical CNN architecture consists of three basic components: convolutional layers, pooling layers, and fully connected layers. The convolution layer is mainly used to extract features from images. The purpose of the pooling layer is to reduce the size of the feature map of the upper layer and

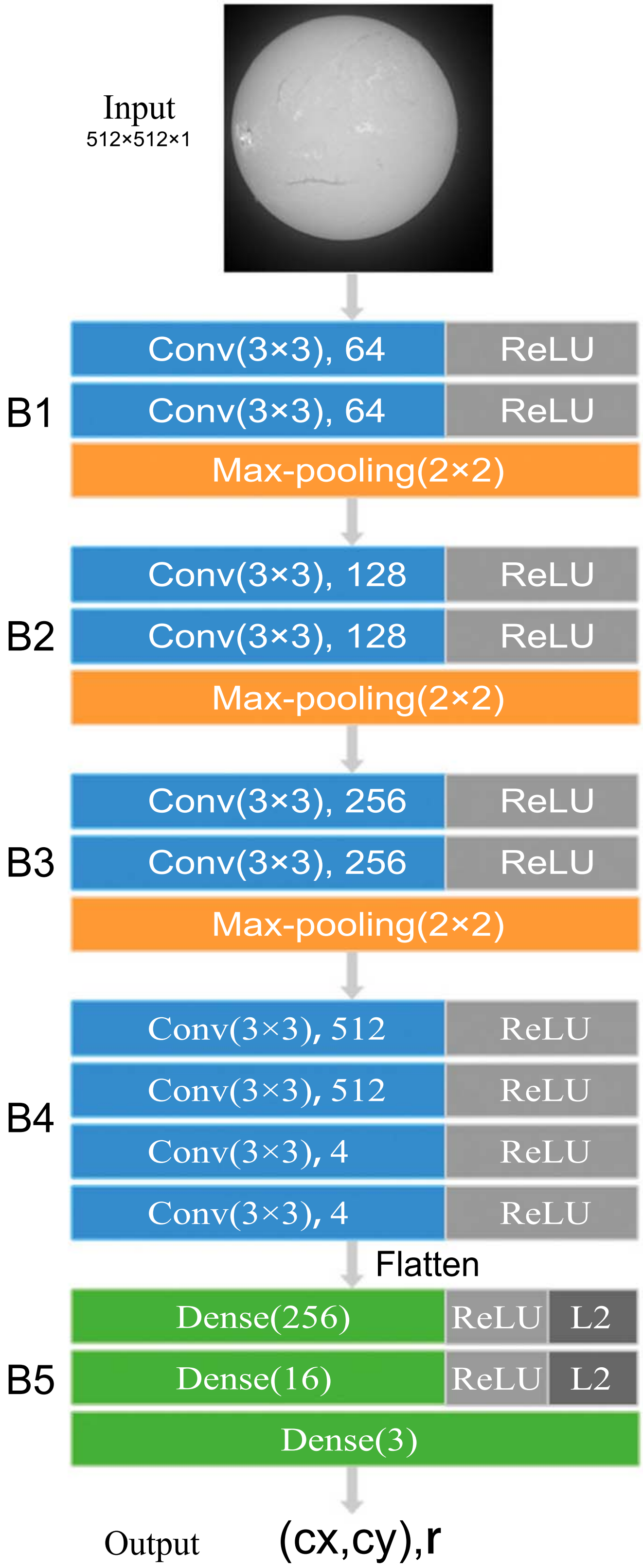

Figure 4. Architecture of the proposed CNNs regression model. The whole model consists of 13 layers with 10 convolutional layers and three fully connected layers. The input is a single-channel image of 512 × 512 and the output is three positive real numbers, which are the center coordinates and radius of the solar disk.

reduces the computational complexity of the network. The fully connected layer is mainly used to fuse the feature information extracted before. Generally, there are two kinds of pooling operations, "average" pooling and "max" pooling. Each value of the latter layer is calculated from the mean value or maximum value of the receptive field (for example, 3 × 3) of the previous layer. An effective CNN usually requires several, even hundreds of convolutional layers and pooling layers. A basic convolutional layer usually obtains a series of feature maps, which are generated by convolving the feature maps of the previous layer with the corresponding convolution kernels. Indispensably, we also need to use an activation function to add nonlinear factors. The jth feature map in the lth convolution layer is denoted as H_j^l. Its expression is

$$H_j^l = f\left(\sum_{i=1}^{N} H_i^{l-1} \otimes w_{ij}^l + b_j^l\right), \tag{1}$$

where w represents convolution kernels (also called filters), b is the bias of the lth layer, and f is the activation function. Usually, the rectified linear unit (Nair & Hinton 2010) is chosen as the activation function. Its expression is $\mathrm{Relu}(x) = \max(0, x)$, which aims to change the value of less than 0 in each feature map to 0. Figure 3 intuitively shows a basic image convolution process. Generally, the initial bias is 0. The convolutional layer mainly obtains feature maps by continuously updating the convolution kernels and bias. The pooling layer is mainly to reduce the size of the image, which can effectively retain key feature information and remove redundant information. The network, as it is translation invariant and rotational invariant, is allowed to learn the feature information of images of different dimensions. The entire training process of a CNNs architecture is to iterate continuously to produce a series of optimal convolution kernels and bias to achieve the objective function (also called loss function).

3.2. *Proposed Network Architecture*

We present a CNNs architecture inspired by VGGNet (Simonyan & Zisserman 2014) to regress the center coordinates and radius of the solar disk from Hα full-disk images. VGGNet was the runner-up of the Image Large Scale Visual Recognition Challenge (ImageNet Challenge) in 2014. VGGNet's network depth varies from 11 to 19 layers, including 3 fully connected layers, and the remaining convolutional layers. So far, it is still an important part for feature extraction in various models.

Since the original VGGNet model cannot be used in our current work directly, we have made some improvements to the VGGNet model, especially VGGNet-13. The proposed CNN models consist of five blocks with a total of 13 layers (Figure 4), in which the first three blocks are the same as VGGNet-13's. Each block in the first three blocks contains two convolutional layers and a max-pooling layer that is used to reduce the dimension of feature maps and reduce the amount of calculation. The first two layers in the first block use 64 × 3 × 3 kernels. The kernel sizes of the convolutional layers in the subsequent blocks are 128 × 3 × 3, 256 × 3 × 3, 512 × 3 × 3, respectively. Unlike VGGNet-13, we use four convolutional layers for the fourth block and no max-pooling layer afterwards. The last two convolutional layers' kernel size is 4 × 3 × 3. The last block is composed of three fully connected layers. The first two fully connected layers have 256 and 16 channels, respectively. The L2 regularization is applied to the kernel weights matrix to handle overfitting, which is

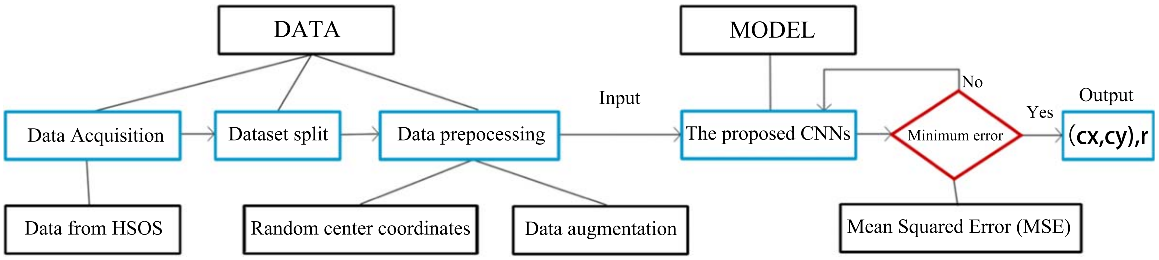

Figure 5. Schematic diagram of the entire task. The whole process consists of two parts: data and model. The blue rectangle represents the task processing order.

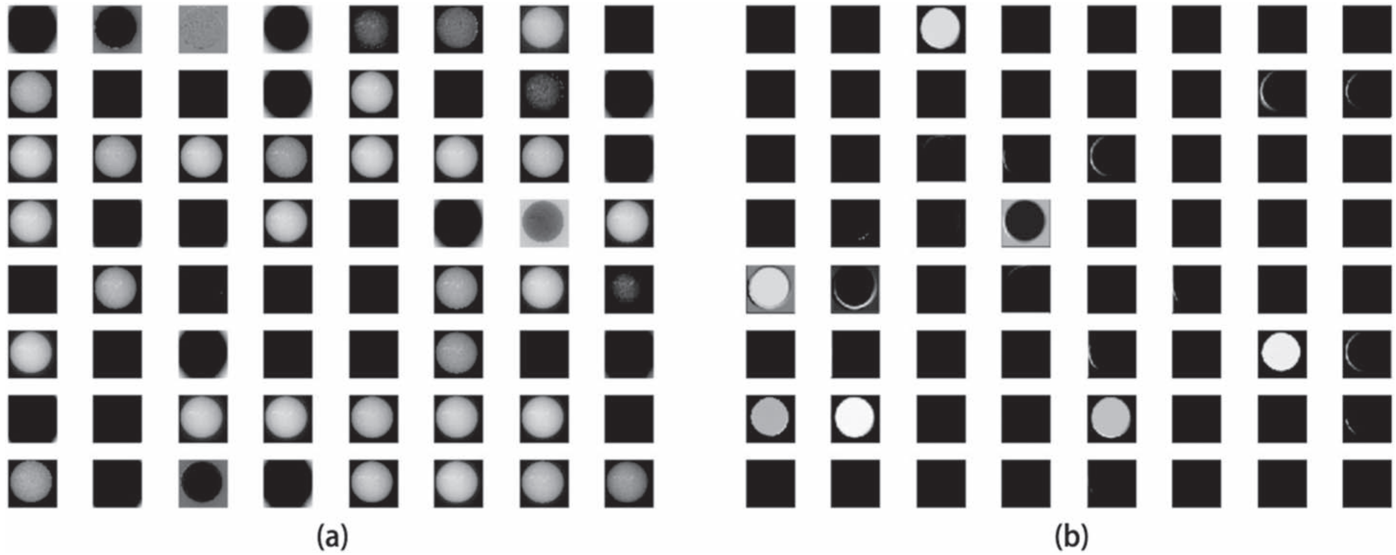

Figure 6. Feature maps of the proposed CNNs model. (a) The feature maps of the convolutional layer in the first block. (b) Part of the feature maps of the convolutional layer in the fourth block.

expressed as

$$H = H_0 + \alpha \sum_{w} w^2, \tag{2}$$

where w is the weight matrix, α is the regularization factor. In the proposed CNN model, α is set to 0.01. Unlike the softmax function used in most CNN models to solve classification problems, the identity function is used in the last fully connected layer of our model. Therefore, the final output is not a probability value, but a positive real matrix of 3×1, which is composed of the center coordinates and the radius of the solar disk. We use single-channel images (grayscale images) as inputs instead of the three-channel color images in the VGGnet. The max-pooling size is 2×2.

There are some other configuration parameters in our network model. In order to avoid the size of the feature map after convolution becoming smaller, we set "padding" equal to "same", which can make the dimensions of input and output of the convolutional layer consistent The convolution kernels are initialized by the He normal initializer (He et al. 2015). The initial bias is 0 tensor.

3.3. Training Details

The entire training process is shown in Figure 5. Before entering the model input, we need to preprocess and enhance the data. However, the data set split is another important thing to be done. Generally, we divide the data set into a training set, a validation set, and a testing set. The training set is mainly used for model learning, which is used to fit the parameters (i.e., weights). The validation set is used to tune the parameters (i.e., hyperparameters, not weights). The testing set is used only to assess the performance of a fully specified model. The whole data set contains 93,686 images, which are divided into three subsets in a ratio of 8:1:1 (Table 1). Furthermore, in order to balance the distribution of the data set, we divide the data set into N/10 parts, where the first eight samples of each part are used as the training set, the ninth is used as the validation set, and the tenth is used as the test set. Due to the hardware limitation of the experiment, the batch size (represents the number of the samples in one iteration of the network) is set as eight. Besides, the samples order is shuffled.

Table 1
Number of Samples in Each Part

Part	Proportion	Quantity
All	1	93686
Training set	0.8	74948
Validation set	0.1	9368
Testing set	0.1	9368

As shown in Figure 6, the entire CNN model needs to travel toward a target (loss function) during training. The loss function of our CNN model is mean squared error (MSE),

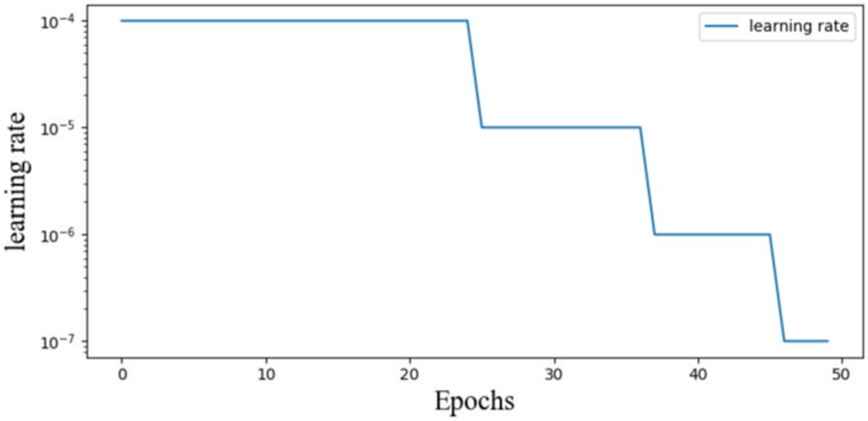

Figure 7. Learning rate curve of one training process.

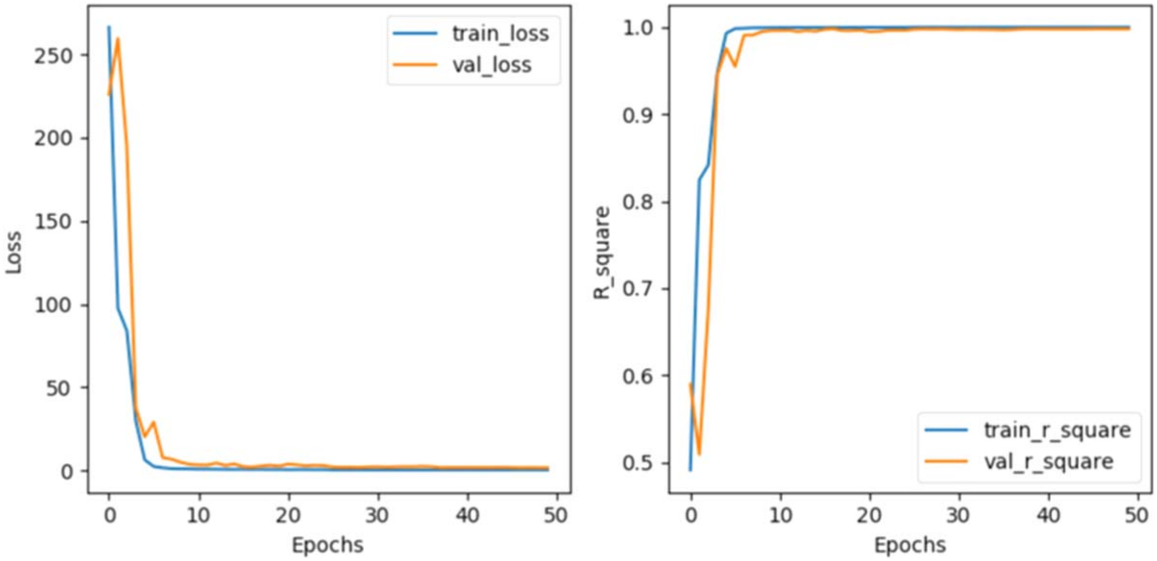

Figure 8. Curves of MSE loss and *R*-square for the training set and validation set. During the training process, the loss of the model on the training set and the validation set can be quickly converged, and the degree of fitting through *R*-square characterization has also rapidly improved.

Table 2
Proportion of Different Relative Errors of 250 Samples

Relative Error	$\delta < 0.01$	$0.01 \leqslant \delta < 0.1$ (0.04)	(0.04)$0.1 \leqslant \delta < 0.2$(0.1)	$\delta \geqslant 0.2$(0.1)
The coordinate of the cx	246 (98.4%)	3 (1.2%)	0 (0%)	1 (0.4%)
The coordinate of the cy	248 (99.2%)	2 (0.8%)	0 (0%)	0 (0%)
Radius	163 (65.2%)	85 (34.0%)	2 (0.8%)	0 (0%)

Table 3
The True Values and the Regressed Results of the Model Under Typical Different Noise Densities

(cx,cy),r	True	0	0.04	0.09
1	(243.19, 241.11), 217.77	(242.76, 241.21), 217.92	(242.87, 241.15), 217.96	(242.76, 241.17), 217.95
2	(232.17, 240.12), 217.24	(231.84, 240.31), 217.69	(231.88, 240.11), 217.69	(231.86, 240.15), 217.66
3	(238.03, 238.17), 216.64	(237.78, 238.54), 217.18	(237.76, 238.42), 217.22	(237.88, 238.48), 217.19
4	(228.54, 246.64), 215.55	(228.18, 246.83), 216.62	(228.20, 246.77), 216.64	(228.25, 246.78), 216.64
5	(240.70, 255.58), 215.37	(240.04, 255.81), 216.46	(240.08, 255.80), 216.53	(240.18, 255.78), 216.51
6	(231.15, 243.96), 211.52	(230.43, 244.78), 214.20	(230.43, 244.71), 214.19	(230.42, 244.75), 214.18
7	(236.26, 243.64), 211.04	(235.30, 244.56), 213.76	(235.22, 244.44), 213.72	(235.25, 244.49), 213.71
8	(235.80, 248.70), 211.76	(235.04, 249.19), 214.28	(235.09, 249.12), 214.34	(235.12, 249.21), 214.34
9	(235.35, 235.77), 213.52	(234.49, 236.49), 215.31	(234.45, 236.47), 215.36	(234.57, 236.40), 215.38
10	(237.81, 237.99), 219.43	(237.67, 238.06), 219.77	(237.55, 238.07), 219.82	(237.72, 238.07), 219.64

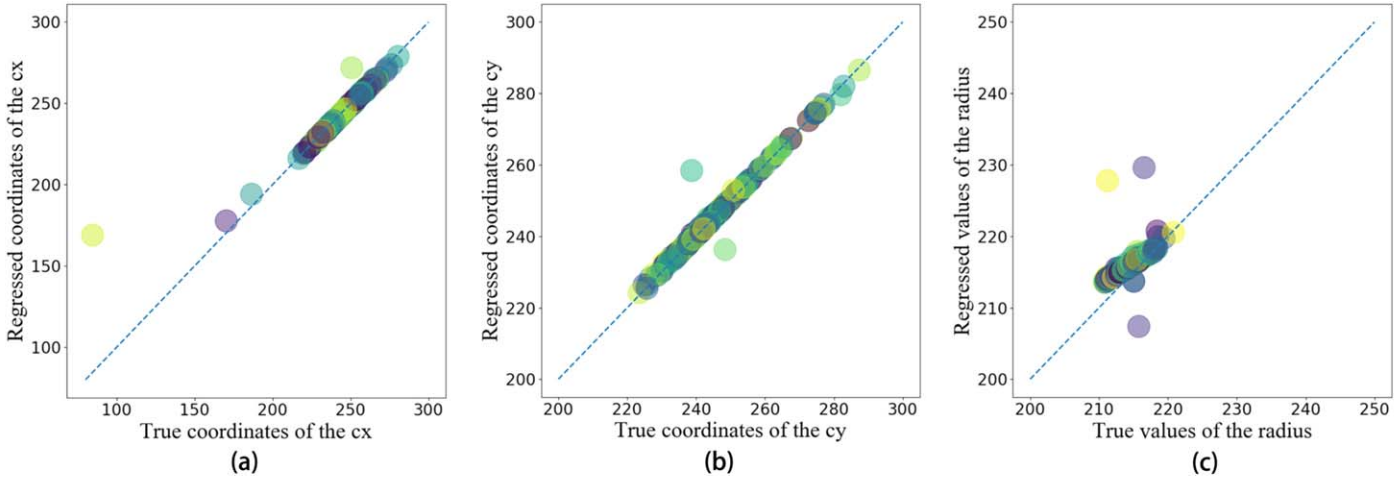

Figure 9. Regressed values and the true values of each image for the center coordinates and the radius of the solar disk. The blue dashed line represents that the regressed values equal the true values. Most samples run through the blue dotted line. There are a few outliers because of cloud pollution, original errors, and incomplete circles.

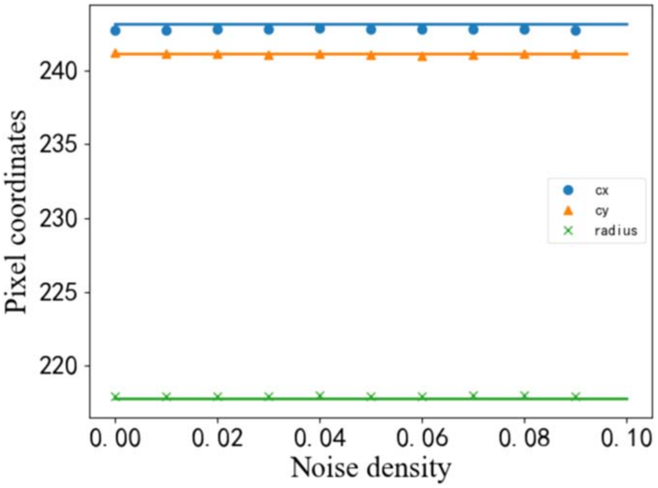

Figure 10. Difference between true values and results under different noise densities. The line represents the true values. The center coordinates and radius of the solar disk maintain stable regression results.

which is given by

$$\mathrm{MSE} = \frac{1}{M}\sum_{m=1}^{M}(y_m - \hat{y}_m)^2, \tag{3}$$

where y_m, $\hat{y}_m$ represent the true value and the predicted value, respectively. The metric of the model is R-square. It is related to the sum of squares of the regression (SSR) and total sum of squares (SST). Its expression is

$$R^2 = \frac{\mathrm{SSR}}{\mathrm{SST}} = \frac{\sum_{m=1}^{M} w_m(\hat{y}_m - \bar{y}_m)^2}{\sum_{m=1}^{M} w_m(y_m - \bar{y}_m)^2} = 1 - \frac{\sum_{m=1}^{M}(\hat{y}_m - y_m)^2}{\sum_{m=1}^{M}(\bar{y}_m - y_m)^2}, \tag{4}$$

where $\hat{y}_m$ is the predicted value, $\bar{y}_m$ is the mean value, and y_m represents the true value. The training process of the model needs MSE as low as possible and R^2 as close to 1 as possible.

We choose "Adam" as the optimizer (Kingma & Ba 2014) with a variable learning rate (it is a hyperparametric value used to adjust the gradient of the loss function) to make the loss as low as possible. The initial learning rate is set to 10^{-4}. If "val_loss" does not decrease after every 8 epochs (epoch: all data in the sample went through one training step), the learning rate will be reduced by $1/10$, and the minimum learning rate is 10^{-8}.

In order to understand the learning process of the entire model, Figure 6 shows the feature maps of the convolution layer in the first block and the fourth block. As can be seen in Figure 6(a), the model mainly extracts the overall information of the images and some features of the image in different contrasts. In Figure 6(b), the model extracts the characteristic of the whitening and ashing of the solar disk. With the deepening of the layers, the model can automatically obtain the essential characteristics of the original solar disk. By continuously adjusting the weights and offsets of all layers, the mapping relationship between features and results is finally established.

Figure 7 shows the learning rate decay curve of our model. Our learning decay strategy is that the val_loss does not decrease after every 8 epochs, and then the learning rate drops by $1/10$. The initial learning rate is 10^{-4}. After about 25 epochs, the first decline begins, and the decline is 10^{-5}. The second drop is approximately after 37 epochs. With the extension of training time and the continuous decline of the learning rate, the model loss and R_ square have no obvious changes, so the training is terminated at that time. The final learning rate stops at 10^{-7}.

4. Experimental Results

The CNNs have a strong ability of self-learning, which can extract features automatically according to the data set provided, and then establish a nonlinear mapping relationship between the problem and the result. In this study, we use the proposed CNN model to regress the center coordinates and the radius of the solar disk from full-disk Hα images. In the left panel of Figure 8, the training loss decreases sharply in the first 6 epochs, and the loss of validation begins to decline sharply in the first 10 epochs. The final convergence of training loss is about 0.28, and the final convergence of validation loss is about

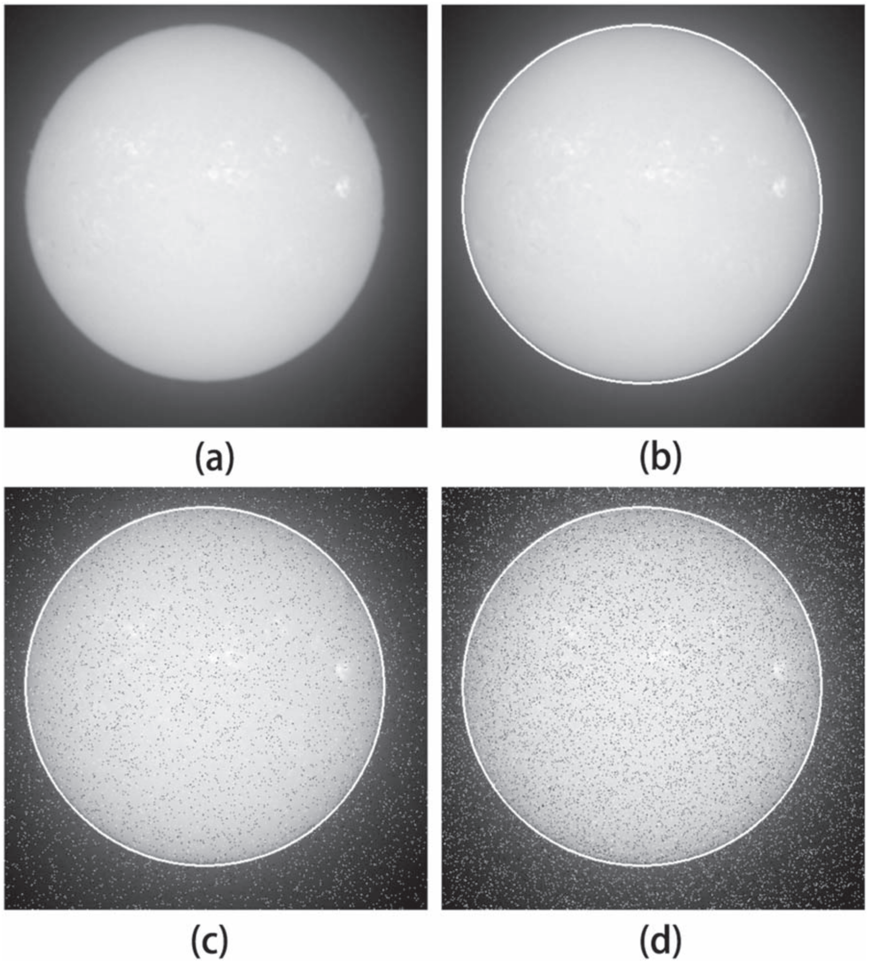

Figure 11. A visualization results of model regression. The white circles represent the results: (a) the original image, (b) a result without noise, (c) a result with a noise density of 0.04, and (d) a result with a noise density of 0.09.

1.6. In the right panel of Figure 8, the *R*-squares of the training and the validation rises sharply in almost corresponding positions, respectively. They all stabilized above 0.99, and the training *R*-square is relatively higher. Furthermore, the average loss of our model on the testing set is 1.75, and the average *R*-square is 0.9974. This shows that our model has good learning ability.

In order to demonstrate that our model is more versatile, we evaluate our model not only on the testing set but also on a independent testing set consisting of 250 images, which is not from the previous data set. In an ideal state, the regressed value should be equal to the true value, as marked by the blue dashed line in Figure 9. The "true" values are obtained by the Hough transform and the least-squares method, and then verified manually. Colored circles represent 250 samples in the new testing set. In Figure 9(a), there are only a few circles outliers, and they are all on the upper left of the blue dotted line. That means the regressed *x*-center (cx) coordinate values are higher than the true values. The vast majority of samples run through the blue dotted line. In Figure 9(b), there are fewer outliers, and they are mainly concentrated along the blue dotted line. That means the regressed *y*-center (cy) coordinate values are larger or smaller than the true values. In Figure 9(c), most of the samples are around the blue dotted line with a few outliers. The following reasons may cause the above outliers: (1) It is inevitable that the uncertainty exists in the data set because of the unsharp edge of the chromosphere, and our model also learns this uncertainty. (2) There are thick cloud pollutions in the outliers, but the model does not learn the features from such clouds. (3) The incomplete solar disk in some of the data causes a relatively larger computational inaccuracy. Probably, data with the latter two reasons are of bad quality and can be removed in the subsequent scientific research.

In addition, we calculate the relative error between the regressed values of the solar disk and the true values,

$$\delta = \frac{|T - P|}{T}, \qquad (5)$$

where *T* is the true value, and *P* represents the predicted value. Table 2 shows the proportion of different relative errors. The relative error below 0.01, for the cx coordinate of the center, accounts for 98.4%. The ratio of relative error between 0.01 and 0.1 is 1.2%. As mentioned above, there are a few outliers with a relative error of less than 1%. For the coordinate of the cy, there is no relative error greater than 0.1. 99.2% of them have a relative error below 0.01. For the radius of the solar disk, the relative error of 65.2% of the samples is controlled below 0.01, and 34% of the samples are between 0.01 and 0.04. Only 0.8% of samples are between 0.04 and 0.1 Although the relative error of the radius does not perform as well as the center coordinates, it still has a relatively low relative error.

Furthermore, it takes an average of 0.953 s to calculate the center coordinates and radius of each image with the CPU of a standard PC. In particular, a standard PC with a 1070Ti GPU only takes an average of 0.024 s to complete the same task.

We also analyze the performance of the trained model under different noise densities. Our model is applied to 10 images with salt-and-pepper noise added, and these images are randomly selected from 250 images. Each image is combined with a noise density of 0–0.1 (excluding right endpoint), whose step size is 0.01. Table 3 shows a comparison between the model regression results and true values under typical different noise densities. By comparison, it is found that under different noise densities, the difference of the center coordinates of the solar disk is still very small, about 1 pixel at the maximum, and the maximum difference of the radius is about 2-3 pixels. Figure 10 depicts the difference between true values and results of the first sample in Table 4. This shows that the proposed model can still regress the expected results stably after part of the solar image is contaminated. Figure 11 is a visualization of its regression results.

5. Discussion and Conclusion

In this paper, we present a new approach to obtain the center coordinates and the radius of the solar disk in full-disk Hα solar images. The deep CNN is used to automatically learn the image features to achieve the regression of the center coordinates and the radius. The data set is available at https://Sun.bao.ac.cn/hsos_data/download/solar-center-radius-data set/, and the source code is open source at https://github.com/GF-Zhu/RCRS. The model uses Hα images as input, and directly obtains three positive real values, the x- and y-center coordinate and the radius of the solar disk. In order to increase the generalization ability of the model, we add a randomization strategy for the center of the solar disk during the preprocessing process. To increase the anti-noise ability of the model, we add a certain amount of salt-and-pepper noises to the training data through data augmentation. We also add L2 regularization in the fully connected layer in the last block to suppress overfitting. The evaluation metric of our model during the training process is R-square. The R-square of the model on the testing set is 0.9974. To better evaluate the model in practical application, we further use an independent testing set consisting of 250 images. The relative error of the regressed center coordinates of the model is mainly controlled below 0.01 and that of the radius is below 0.04. This shows that our approach is a viable and efficient strategy for the regression of the center coordinates and the radius of the solar disk.

Our proposed CNN model has a certain ability to resist noise. Moreover, it can add more strategies to the data augmentation to improve the model. For the least square method, we need to eliminate noise as much as possible before using it. The Hough transform circle detection is not only computationally intensive but also needs to set some parameters in advance, such as the radius range. Our method can automatically extract features from the data set without prior knowledge. It avoids the shortcomings of the other two methods. It should be noted that our method can also be used for real time applications when the trained model is available. And in addition, the CNN model has the promising ability to detect the image quality and screen out bad-quality data especially for the ground-based observations. This method can be provided as class library or a third-party tool after continuous improvement and be of future interest.

The work was funded by National Science Foundation of China Grant Nos. u1531247 and 11427901, the 13th Five-year Informatization Plan of Chinese Academy of Sciences, Grant No. XXH13505-04 and the special foundation work of the Ministry of Science and Technology of the China Grant No. 2014fy120300. The authors thank the referee for valuable suggestions and recommendations.

ORCID iDs

Gaofei Zhu https://orcid.org/0000-0002-0405-7018
Dongguang Wang https://orcid.org/0000-0003-4344-4093
Xiao Yang https://orcid.org/0000-0003-1675-1995

References

Ai, G. X., & Hu, Y. 1986, BeiO, 8, 1
Ballard, D. H. 1981, Pattern Recognit., 13, 111
Cabrera-Vives, G., Reyes, I., Förster, F., Estévez, P. A., & Maureira, J.-C. 2017, ApJ, 836, 97
Curto, J. J., Blanca, M., & Martínez, E. 2008, SoPh, 250, 411
Deng, Y., Ai, G., Wang, J., et al. 1997, SoPh, 173, 207
Goel, S., & Mathew, S. K. 2014, SoPh, 289, 1413
Gonzalez, R. C., & Woods, R. E. 2007, Digital Image Processing (New York: Pearson Education)
Hao, Q., Fang, C., Cao, W., & Chen, P. 2015, ApJS, 221, 33
Harvey, J. W., Hill, F., Hubbard, R. P., et al. 1996, Sci, 272, 1284
He, K., Zhang, X., Ren, S., & Sun, J. 2015, in 2015 IEEE Int. Conf. on Computer Vision (ICCV) (New York: IEEE), 1026
Hill, F., Fischer, G., Forgach, S., et al. 1994, SoPh, 152, 351
Ho, M., Rau, M. M., Ntampaka, M., et al. 2019, ApJ, 887, 25
Kingma, D., & Ba, J. 2014, arXiv:1412.6980
Lanusse, F., Ma, Q., Li, N., et al. 2018, MNRAS, 473, 3895
Lecun, Y., Bottou, L., Bengio, Y., et al. 1998, Proc. IEEE, 86, 2278
Li, X., Zheng, Y., Wang, X., & Wang, L. 2020, ApJ, 891, 10
Lin, G., Zhu, G., Yang, X., et al. 2020, ApJS, 249, 11
Liu, S. 2020, AdSpR, 65, 1054
Nair, V., & Hinton, G. E. 2010, in Proc. 27th Int. Conf. on Machine Learning (ICML 2010) (Madison, WI: Omnipress), 807, https://icml.cc/Conferences/2010/papers/432.pdf
Park, E., Moon, Y.-J., Shin, S., et al. 2018, ApJ, 869, 91
Pötzi, W., Veronig, A. M., Riegler, G., et al. 2015, SoPh, 290, 951
Simonyan, K., & Zisserman, A. 2014, arXiv:1409.1556
Van Oort, C. M., Xu, D., Offner, S. S., & Gutermuth, R. A. 2019, ApJ, 880, 83
Veronig, A., Steinegger, M., Otruba, W., et al. 2000, in ESA SP-463, The Solar Cycle and Terrestrial Climate, Solar and Space Weather, ed. A. Wilson (Noordwijk: ESA), 455
Watson, F., Fletcher, L., Dalla, S., & Marshall, S. 2009, SoPh, 260, 5
Yang, M., Tian, Y., Liu, Y., & Rao, C. 2018, SoPh, 293, 81
Yuan, Y., Shih, F., Jing, J., Wang, H., & Chae, J. 2011, SoPh, 272, 101
Zhang, H.-Q., Wang, D.-G., Deng, Y.-Y., et al. 2007, ChJAA, 7, 281
Zhao, C., Lin, G., Deng, Y., & Yang, X. 2016, PASA, 33, e018
Zharkov, S., Zharkova, V. V., & Ipson, S. S. 2005, SoPh, 228, 377
Zheng, Y., Li, X., & Wang, X. 2019, ApJ, 885, 73
Zhu, G., Lin, G., Wang, D., Liu, S., & Yang, X. 2019, SoPh, 294, 117