Research Article

Research on Mount Wilson Magnetic Classification Based on Deep Learning

Yuanbo He,[1] Yunfei Yang [ID],[1,2] Xianyong Bai,[2] Song Feng,[1] Bo Liang,[1] and Wei Dai[1]

[1]*Faculty of Information Engineering and Automation, Yunnan Key Laboratory of Computer Technology Application, Kunming University of Science and Technology, Kunming 650500, Yunnan, China*
[2]*CAS Key Laboratory of Solar Activity, National Astronomical Observatories, Beijing 100012, China*

Correspondence should be addressed to Yunfei Yang; yangyf@kust.edu.cn

The Mount Wilson magnetic classification of sunspot groups is thought to be meaningful to forecast flares' eruptions. In this paper, we adopt a deep learning method, CornerNet-Saccade, to perform the Mount Wilson magnetic classification of sunspot groups. It includes three stages, generating object locations, detecting objects, and merging detections. The key technologies consist of the backbone as Hourglass-54, the attention mechanism, and the key points' mechanism including the top-left corners and the bottom-right corners of the object by corner pooling layers. These technologies improve the efficiency of detecting the objects without sacrificing accuracy. A dataset is built by a total of 2486 composited images which are composited with the continuum images and the corresponding magnetograms from HMI and MDI. After training the network, the sunspot groups in a composited solar full image are detected and classified in 3 seconds on average. The test results show that this method has a good performance, with the accuracy, precision, recall, and mAP as 0.94, 0.93, 0.94, and 0.90, respectively. Moreover, the flare productivities of different types of sunspot groups from 2011 to 2020 are calculated. As $I_{tot} \geq 1$, the flare productivities of $\alpha, \beta, \beta\gamma, \beta\delta$, and $\beta\gamma\delta$ sunspot groups are 0.14, 0.28, 0.61, 0.71, and 0.87, respectively. As $I_{tot} \geq 10$, the flare productivities are 0.02, 0.07, 0.27, 0.45, and 0.65, respectively. It means that the $\beta\gamma, \beta\delta$, and $\beta\gamma\delta$ types are indeed very closely related to the eruption of solar flares, especially the $\beta\gamma\delta$ type. Based on the reliability of this method, the sunspot groups of the HMI solar full images from 2011 to 2020 are detected and classified, and the detailed data are shared on the website (https://61.166.157.71/MWMCSG.html).

1. Introduction

Sunspots are the typical manifestation of the strong magnetic fields on the solar surface. They are closely associated with solar activities, such as solar flares and coronal mass ejections [1–8]. These activities will disrupt the atmosphere of space and Earth, affect the ground short wave radio communication, and produce hazards such as magnetic storm [5].

Sunspots tend to appear in magnetically bipolar groups. Even a unipolar spot group really has dual polarity; the magnetic field strength of the other polarity is not intense enough to cause a visible spot. With the evolving of sunspot groups, they show various morphologies and complex magnetic polarities. Therefore, some classification schemes were proposed to describe the generality of sunspot groups.

Typically, three classification schemes have been successively proposed: Zurich, McIntosh, and Mount Wilson magnetic classification. The Zurich sunspot classification was proposed to classify the sunspots into nine classes, *A* to *J*, comprising almost all stages of sunspots occurring mainly [9, 10]. Later, the Zurich classification was modified and expanded to the McIntosh classification, by adding size, stability, and complexity that appeared to correlate with flares [11, 12]. The form was expanded as Zpc, where Z is the modified Zurich class, p is the type of principal spot, and c is the degree of compactness in the interior of the group. They present that the correlations with flares are excel with earlier Zurich classification, especially the larger flares correlating with types Dki and Eki [12]. Both the Zurich classification and the McIntosh classification mainly depend on the morphology of sunspots, which require only white-light

observations. The Mount Wilson magnetic classification was proposed earlier [13]. The scheme classifies sunspot groups into eight classes based on their morphological and magnetic properties together (for the detailed please, see Section 2). Therefore, both white-light observations and magnetograms are needed. Most M-class and X-class flares are found to be erupting above the complex sunspot groups, just like the $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ sunspots groups [14]. In particular, the $\beta\gamma\delta$ sunspot groups have very high probabilities of flares [15]. Whichever classification schemes indicate that the more complex the morphological structure and the magnetic polarity of a sunspot group is, the higher the probability of the flare is [16, 17].

Previous authors have presented some methods for the above three classification schemes. For instance, Nguyen et al. [18] used machine learning techniques, including decision trees, rough sets, hierarchical clustering, and layered learning methods, to classify sunspot groups into seven-class modified Zurich classes. Abd et al. [19] employed support vector machines to achieve modified Zurich classification. Colak and Qahwaji [20] adopted the traditional image processing algorithm to detect sunspots, such as morphological operator, thresholding, and region growing; the features of sunspot groups, like length, height, and area, were then extracted; the McIntosh classification was determined finally by a decision tree using the extracted features. Hong et al. [21] performed the Mount Wilson magnetic classification according to features such as plus polarity, minus polarity, and magnetic neutral line, which are extracted from the continuum images and magnetograms after doing morphological operations and threshold methods. Padinhatteeri et al. [22] proposed an algorithm called SMART-DF to detect the δ type. They also adopted the threshold method to identify the umbra and penumbra, and then the pairs with a distance less than 2° are reserved, where the distance is between each possible pairing of opposite-polarity umbrae. Finally, the pairs of opposite-polarity spots that pass several conditions involving the areas of umbra and penumbra are marked as δ type. Recently, Fang *et al.* [23] used a deep learning method based on Lenet-5 [24, 25] to classify the sunspot groups into α, β, and β-x types, where the sunspot type falls into three categories based on the Mount Wilson classification. However, the automatic method for the whole Mount Wilson classification in sunspot groups is still lacking.

In recent years, deep learning methods use various machine learning algorithms based on multilayer neural networks to solve problems [26], such as image detection and classification [27, 28]. The core of these methods is deep feature learning, which acquires hierarchical feature information through hierarchical networks. CornerNet-Saccade [29] is a new object detection algorithm, whose detection scheme based on the attention mechanism [30] and the key points [31–33] make it have advanced performance in the fields of object detection. This paper proposes a method for the Mount Wilson magnetic classification of sunspot groups based on CornerNet-Saccade. It obtains a good performance with the accuracy, precision, recall, and mAP, which all are above 0.90.

The main organizational structure of this paper is as follows. Section 2 introduces the Mount Wilson magnetic classification scheme briefly. Section 3 describes the data source and how to build the dataset. In Section 4, the method including the main steps is listed. Section 5 details the experiments, results, and evaluation of the method. Sections 6 and 7 discuss and summarize the results, respectively.

2. Mount Wilson Magnetic Classification of Sunspot Groups

The Mount Wilson Observatory regards the bipolar sunspot group as the basic type and other types as the deformations of the bipolar sunspot group according to the polarities of the magnetic fields [13]. Following the classification scheme, α is a unipolar sunspot group, and β is a distinct bipolar sunspot group with opposite polarities. γ is a complex sunspot group with irregular polarities. δ is a sunspot group with umbra which have opposite polarities and are separated by less than 2° within one penumbra. $\beta\gamma$ is a bipolar sunspot group with more than one continuous line. Besides that, if the sunspot groups contain one or more δ spots, δ spots shall be added to the corresponding types. This classification is also called Hale class. The detailed description of the classification is as follows:

(i) α is a unipolar sunspot group

(ii) β is a bipolar sunspot group, with the simple and distinct division between opposite polarities

(iii) γ is a complex sunspot group with irregular polarities

(iv) δ is a sunspot group with umbra having opposite polarities within a penumbra and spans less then 2°

(v) $\gamma\delta$ is a γ sunspot group containing one or more δ sunspots

(vi) $\beta\gamma$ is a bipolar sunspot group with more than one continuous polarity reversal line

(vii) $\beta\delta$ is a β sunspot group containing one or more δ sunspots

(viii) $\beta\gamma\delta$ is a $\beta\gamma$ sunspot group containing one or more δ sunspots

3. Data

The data used in this paper come from the Helioseismic and Magnetic Imager (HMI) of the *Solar Dynamics Observatory (SDO)* [34] satellite and Michelson Doppler Imager (MDI) of the *Solar and Heliospheric Observatory (SOHO)* [35]. We used the continuum images and magnetograms, where the HMI data span from June 2010 to April 2020, and the MDI data span from July 2000 to September 2001. Both of them are selected with 12-hour interval.

Since the Mount Wilson magnetic classification needs to take into account the morphological structures and magnetic properties of the sunspot groups, we first composited the sunspot morphological structure information from the continuum image and the magnetic properties information

from the magnetogram into one image. The main steps are as follows (Figure 1):

(1) The continuum image and the cotemporal magnetogram are smoothed and normalized separately.

(2) The penumbra boundaries and umbra regions of the continuum image are extracted by the threshold method [22]. They are represented by A and B, respectively. The positive and negative polar regions of the magnetogram are extracted, which are represented by C and D, respectively.

(3) $E = B \cap C$ and $F = B \cap D$, where E is the positive region of the umbra, and F is the negative region. $\cap$ is the intersection operation.

(4) The penumbra boundaries corresponding to A in the continuum image are contoured with green; the positive regions of the umbra corresponding to E are filled with red; and the negative regions of the umbra corresponding to F are filled with blue. Figure 1 shows an example. Figure 1(a) is the continuum image, Figure 1(b) is the magnetogram, and Figure 1(c) is the composited image.

A total of 2486 composited images were used to build the Mount Wilson magnetic classification datasets. The training set is used to train the deep learning network model, which consists of 1886 composited images. The test set is used to test the performance of the model, which consists of 600 composited images. The HMI data observed daily at 08:00:00 UT and 20:00:00 UT from 1 June 2010 to 31 December 2016 were mainly used for building the training set, and the data from 1 January 2017 to 30 April 2020 were used for the test set. All the samples were selected from them representatively. The MDI data which were observed in July 2000 (NOAA 9601) and September 2001 (NOAA 9087) were supplements because there were no γ and $\gamma\delta$ types occurring from 2010 until April 2020 according to the Solar Region Summary (SRS) text file (https://www.swpc.noaa.gov/). The γ and $\gamma\delta$ types only appeared in NOAA 9601 and NOAA 9087 from 1996 until April 2020. Therefore, all the scarce data including γ and $\gamma\delta$ types were added to the training set and the test set without intersection. Besides that, some data including the $\beta\delta$ and $\beta\gamma\delta$ types from 2010 to 2016 were moved from the training set to the test set because these two types have been relatively lacking since 2017. It should be noted that, from 1996 until now, the SRS file has not reported δ type. Therefore, the remainder of this paper no longer involves δ type.

LabelMe tool [36] was used to label the samples of the dataset. Each sample, a sunspot group, was given a label according to the combination of the definition of Mount Wilson magnetic classification [13] and the label of the SRS file. If the label of a sunspot group from the SRS was doubtful, then the label was mainly decided by the definition of Mount Wilson magnetic classification. A total of 10,286 samples were labeled in the datasets, including 7805 samples in the training set and 2433 samples in the testing set. Table 1 lists the numbers and proportions of all samples for various types of sunspot groups in the datasets. It can be seen that the

numbers of samples differ greatly, which is due to the quite difference in occurrences of different types of sunspot groups. For example, the β type appears frequently, while the γ type appears rarely. We had tried to solve the problem of the unbalanced numbers of the samples by oversampling. However, the experimental results show that the classification effect is not significantly improved. Therefore, this paper does not do anything to balance the samples of various types of sunspot groups.

4. Method

We adopted a new deep learning model, CornerNet-Saccade, to implement the Mount Wilson magnetic classification of sunspot groups. It mainly includes three stages, generating object locations, detecting objects, and merging detections. The main flow of the method is shown in Figure 2. The stage of generating object locations is plotted in blue, which is to find approximate locations and rough sizes of the sunspot groups on the solar full images. The stage of detecting objects is plotted in orange, which is to further determine the specific locations of the sunspot groups on the basis of the first stage and give the classification and confidence score of the sunspot groups. The third stage of merging detections is plotted in red, which is to merge the detections for eliminating the redundant boxes.

The deep learning method needs training the network by the samples in the training set first, which is to adjust the parameters of networks by executing the feedforward propagation and back propagation in iterations. The propagation algorithm performs a series of operations such as convolution and pooling to obtain the classification probability and bounding box of the objects. The loss value of the loss function, which is used to evaluate the error between the estimated value and the true value, is then fed back to the entire network by the back propagation algorithm; that is, the gradient of each layer of the networks is calculated using the gradient descent algorithm to continuously adjust the weight of each parameter to minimize the loss. The training procedure is iterated until the loss value converges steadily. The main steps, seen in Figure 2, are detailed in the following (Figure 3).

(1) Downsizing the composited images for reducing inference time under limited memory.

(2) Inputting the downsized images to the backbone network, Hourglass-54 [37]. The Hourglass-54 network is composed of three hourglass modules, which has a total depth of 54 layers. Each hourglass is a modular network with a symmetrical structure (see Figure 3). The hourglass network first applies three stages of convolution and downsampling layers to reduce the size of the input feature maps and then upsamples the features back to the original resolution by three stages of convolution and upsampling layers, combining features across scales. Downsampling is achieved by convolution with stride of 2, and upsampling is achieved by nearest neighbour interpolate [38]. One residual unit [39] is applied

(a)   (b)   (c)
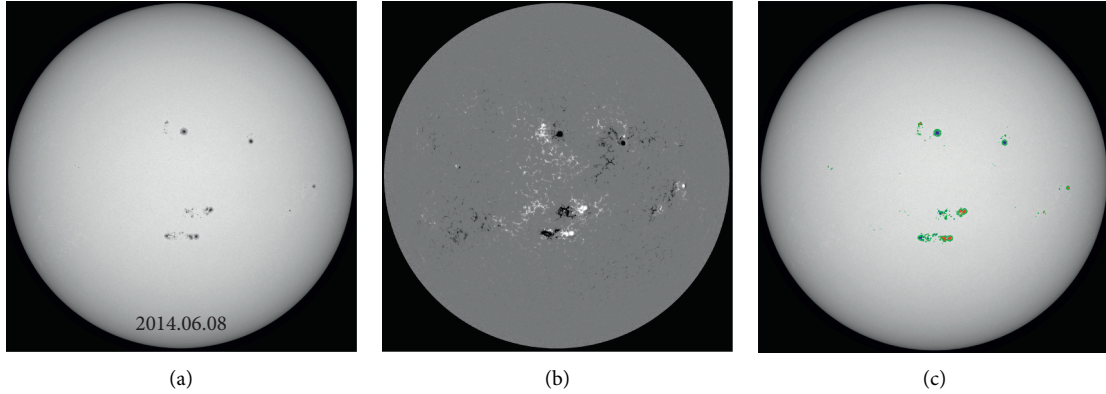
FIGURE 1: An example of compositing the morphological structure information of sunspots from the continuum image and the magnetic properties' information from the magnetogram. The time is 20:00:00 UT on 8 June 2014. (a) The continuum image. (b) The magnetogram. (c) The composited image, in which the penumbra boundaries of the sunspot groups are contoured with green, and the positive umbra and the negative umbra are filled with red and blue, respectively.

TABLE 1: The numbers and proportions of the samples of various types.

Type	α	β	γ	$\gamma\delta$	$\beta\gamma$	$\beta\delta$	$\beta\gamma\delta$
Number	2953	5086	13	28	1470	135	601
Proportion (%)	28.71	49.45	0.13	0.27	14.29	1.31	5.84

after each downsampling layer and upsampling layers, which corresponds to each box in Figure 3. Each residual unit is composed of two parts, three weight layers, and one identity mapping (or shortcut connections). Each weight layer is used to obtain deep-level information, which is designed as convolutions and following ReLU (Conv-ReLU). The identity mapping is used to retain the original information, which is designed as a 1×1 convolution for matching dimensions with the output of the weight layer. There are ten residual units in the hourglass module, which better solves the vanishing degradation problem of deep neural networks and improves the network performance. A total of three hourglass modules together allow for repeated bottom-up, top-down inference across scales, which capture and consolidate information across all scales of the image. The outputs of Hourglass-54 are a set of feature maps.

(3) Applying 3×3 Conv-ReLU module and a 1×1 Conv-Sigmoid module on each feature map to predict attention maps by the attention mechanism. The attention mechanism is derived from the saccades mechanism, which refers to a sequence of rapid eye movements to fixate different image regions [40].The feature maps at finer scales are used for smaller objects and the ones at coarser scales are for larger objects. A total of three size scales of attention maps are predicted, corresponding to small, medium, and large objects, respectively. The attention map indicates the approximate locations and rough sizes of the object. Meanwhile, Hourglass-54 also

generates bounding boxes for the detected objects in the downsized image. The bounding boxes are obtained by detecting the two key points, the top-left corners and the bottom-right corners of the object, through corner pooling layers [33]. If the Intersection over Union (IoU, Rezatofighi et al. [41]), the ratio of intersection and union of the bounding box and ground truth, is greater than a threshold (set as 0.3 in this work), the bounding box is reserved because it is more likely to contain the targets. So far, the approximate locations and sizes of the objects from the attention maps and bounding boxes are obtained separately.

(4) the downsampled image at each possible location to find the locations of the objects more accurately. For the locations obtained from the attention maps, the zoom ratios of small, medium, and large objects are commonly set as 4, 2, and 1, respectively. For the locations obtained from the bounding boxes, the downsized image is enlarged according to the size of the bounding box.

(5) the enlarged image back to the original image and then cropping the regions by taking the locations as the center points.

(6) the regions according to the scores and then picking up the top k locations with the highest scores.

(7) the possible objects in each region through the second Hourglass-54. The bounding boxes of the objects are fine adjusted by detecting the two key points, the top-left corners and the bottom-right corners of the object, through the corner pooling layers.

(8) Training the network. The loss function of the model consists of four losses with different weights [33]: a variant of focal loss, L_{det} [42], the smooth L1 loss at ground-truth corner locations, L_{off} [43], the loss of grouping the corners, L_{pull} [44], and the loss of separating the corners, L_{push} [44].
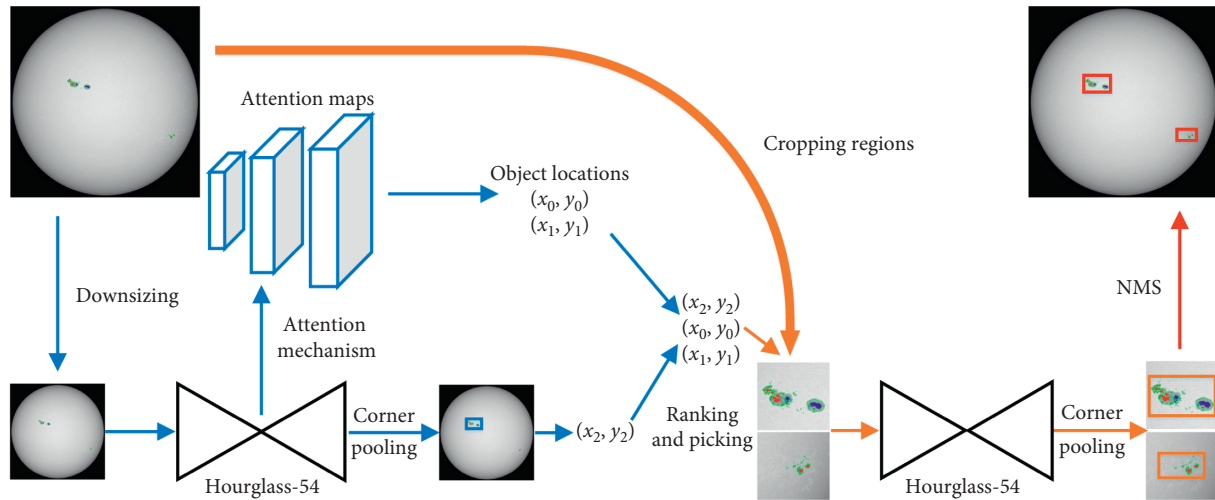
FIGURE 2: The main flow of the network. The first stage of generating object locations is plotted in blue, which is to find approximate locations and rough sizes of the sunspot groups on the solar full images. The second stage of detecting objects is plotted in orange, which is to further determine the specific locations of the sunspot groups on the basis of the first stage and give the classification and confidence score of the sunspot groups. The third stage of merging detection is plotted in red, which is to merge the detections for eliminating the redundant boxes.
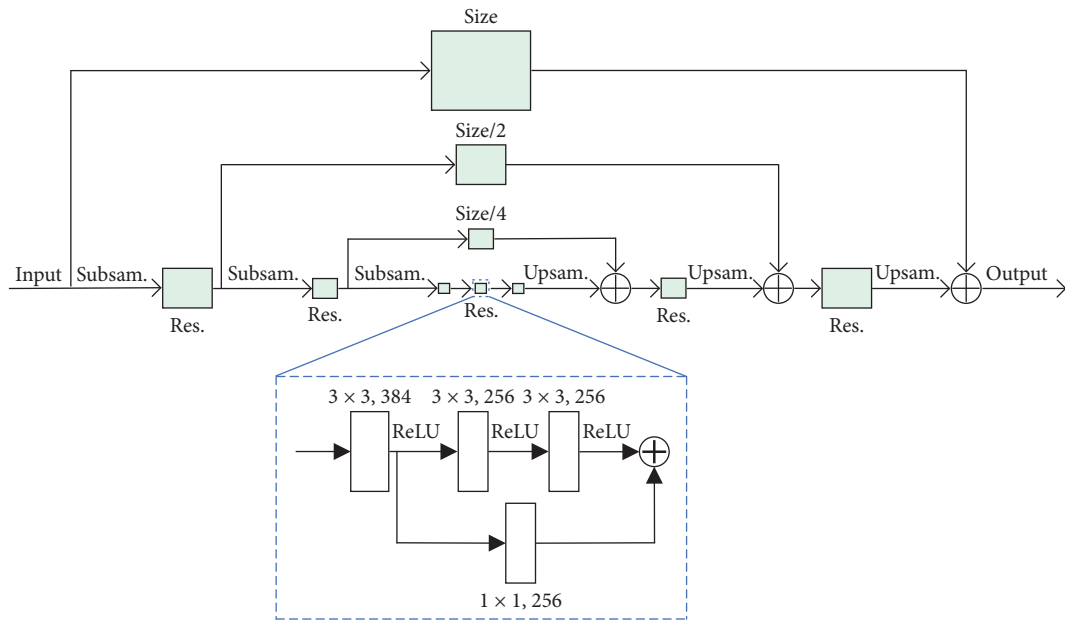

FIGURE 3: An illustration of a single "hourglass" module. Each hourglass is a modular network with a symmetrical structure. Each box represents a single residual module.

(9) Testing the network. The feedforward propagation is performed only once with the flow of Figure 2, except for back propagation with loss functions. Besides that, all detected bounding boxes are merged in the solar full image, where the redundant bounding boxes are eliminated based on Nonmaximum Suppression (NMS) [45]. The typical NMS is to find local maximum value and suppress nonlocal maximum values for each class, which aims to eliminate the redundant boxes of each class. This work modifies the typical NMS as eliminating the redundant boxes for all bounding boxes as a whole, not repeating for each class. On average, it takes about three seconds to process one image.

The experiments were performed on a personal computer equipped with a GeForce RTX$^{\text{TM}}$ 2080 graphics card. The main programs deployed were Python 3.7, PyTorch 1.2, CUDA 10.1, GCC 7.4, etc. Adam [46] was chosen to optimize both losses for the attention maps and object detections. After repeated experiments, the hyperparameters such as the learning_rate, decay_rate, batch_size, and IoU of the network model were set to 0.000025, 10, 8, and 0.01, respectively. The loss value of the network converges stably after 335,000 iterations of training, which takes about 110 hours.

5. Results

5.1. Instance. Figure 4 shows two cases. The sunspot groups with solid line boxes are detected by our method, where the classification results are displayed at the upper left of each box. The active region (AR) number and classification of the sunspot groups given by NOAA are manually labeled to the box below. Note that the reports of NOAA are issued on today using data from yesterday. We marked them by using the correct reports.

It can be seen in Figure 4(a) that our method detects five sunspot groups. A total of three sunspot groups are marked by NOAA: NOAA 12299, NOAA 12297, and NOAA 12298. Our classification results are consistent with them, where the three sunspot groups are classified into β, $\beta\gamma\delta$, and β, respectively. Besides that, two other sunspot groups (red arrows 1 and 2) are detected by our method, and both of them are classified into β type.

In Figure 4(b), eight sunspot groups are detected by our method, while seven are marked by NOAA. There are six sunspot groups that are the same by NOAA and by our method. NOAA 11971 (red arrow 3) is classified as β type by our method; however, it is classified as α type by NOAA. By analyzing the continuum image and the corresponding magnetogram, this sunspot group is a bipolar group, where there is no mixed polarity and no umbra of opposite polarity in the same penumbra region. Therefore, the sunspot group ought to be β type. Besides that, there is a visible sunspot group above NOAA 11973 (red arrow 4). Our method detects the sunspot group and classifies it into β type correctly.

5.2. Evaluation Metrics. We evaluated the performance of our method by four metrics: accuracy, precision, recall, and AP [47]. The definitions are as follows:

$$
\begin{aligned}
\text{accuracy} &= \frac{\text{TP}}{N}, \\
\text{precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\
\text{recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\
\text{AP} &= \int_0^1 P(r)\,\mathrm{d}r,
\end{aligned}
\tag{1}
$$

where N is the total number of samples. The true positive, TP, is the number of items correctly labeled as belonging to the positive class. The false positive, FP, is the number of items incorrectly labeled as belonging to the class. The false negative, FN, is the number of items which are not labeled as belonging to the positive class but should have been. The average precision, AP, is a typical performance metric of combination of precision and recall, which is calculated as the precision averaged across all recall values between 0 and 1.

Table 2 lists the metrics of the test set. Note that γ and $\gamma\delta$ are not taken into account in Table 2, because there are few γ and $\gamma\delta$ types of sunspot groups in the training set and the test set. It is meaningless to evaluate their corresponding performances. The mean accuracy, precision, recall, and mean AP (mAP) are all above 0.90. There are three types in which the AP values reach up to 0.90, namely, β, $\beta\gamma$, and $\beta\gamma\delta$. On the other hand, the APs of α and $\beta\delta$ types are a little lower than 0.90. We found that some small sunspots belonging to α type are easily confused with the background, which results in false detections and missed detections. The $\beta\delta$ type has some characteristics confused with β, $\beta\gamma$, and $\beta\gamma\delta$, so there are a few false detections.

5.3. Statistics. Based on the satisfactory results of this method, the solar full images of HMI from January 2011 to April 2020 were fed to the trained network for the Mount Wilson magnetic classification of sunspot groups. A total of 3086 composited images are available; most of them are at 20:00:00 UT and a few of them are near 20:00:00 UT. As a result, a total of 11, 681 sunspot groups are detected and classified. Note that only five types of sunspot groups appeared in this decade. They are α, β, $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$, where the numbers are 3951, 6094, 1158, 110, and 368, respectively. Therefore, only these five types are analyzed in this section. Detailed data are shared on the website (https://61.166.157. 71/MWMCSG.html).

We summarized the total number for each type of sunspot group in months in Figure 5. The x-axis represents time in months, and the y-axis represents the number of sunspot groups. Because no data are provided in some days, we fill the missed data by averaging the previous day and the next day in order to plot the lines continuously in the figure.

The number of sunspot groups belonging to the same type varies greatly over time. For α type, the highest number is 106 in December 2014; the lowest number falls to 1 in February 2020. The difference between the maximum value and minimum value reaches up to 105. For β type, there are two months with the highest number of 144, including November 2011 and May 2013. On the other hand, the lowest number is 0 in October 2019. Besides that, the other types are relatively rare. The highest numbers of $\beta\gamma$ and $\beta\delta$ are 49 and 12, respectively, which occur in February 2014 and November 2014, respectively. For $\beta\gamma\delta$ type, there are two months with the highest number of 18: July 2012 and October 2013. For $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$, there are 24, 63, and 54 months with the lowest numbers of 0, respectively.
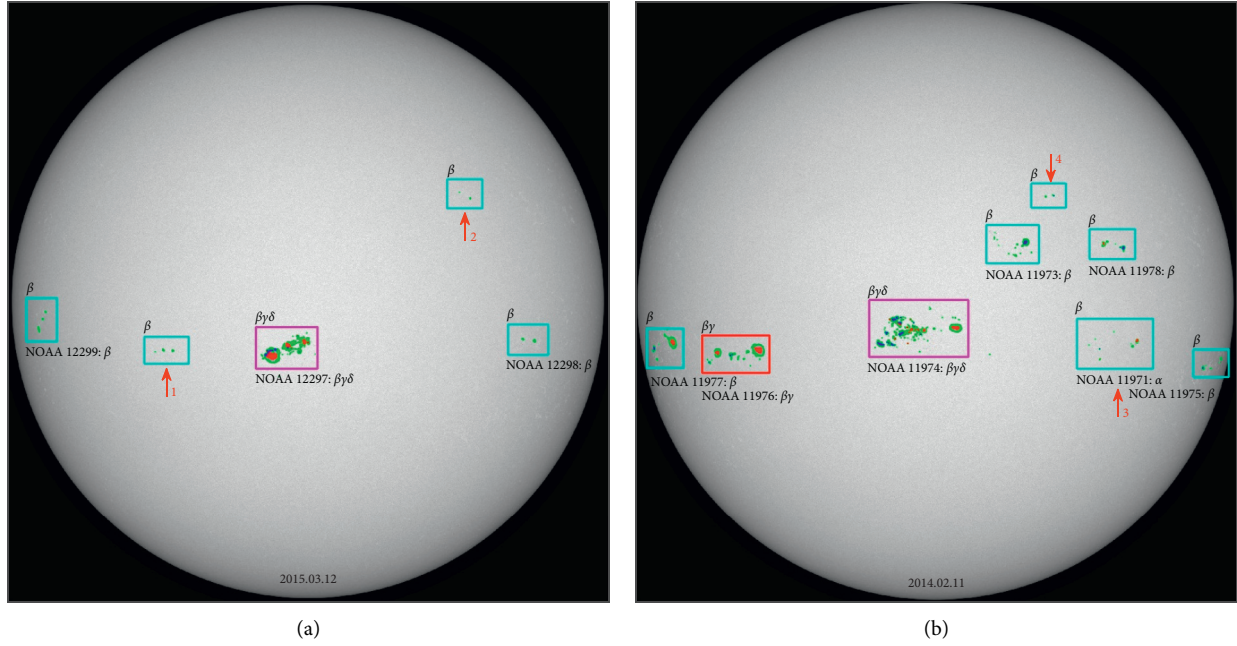
(a)

(b)

FIGURE 4: Two cases. The sunspot groups with solid-line boxes are detected by our method, where the classification results are displayed at the upper left of each box. The AR number and classification given by NOAA are manually labeled at the box below. The times of (a) and (b) are 20:00:00 UT on 12 March 2015 and 20:00:00 UT on 11 February 2014, respectively.

TABLE 2: The metrics of the test set.

Type	Accuracy	Recall	Precision	AP
α	0.90	0.92	0.90	0.89
β	0.95	0.94	0.95	0.93
$\beta\gamma$	0.94	0.92	0.94	0.90
$\beta\delta$	0.93	0.91	0.93	0.87
$\beta\gamma\delta$	0.96	0.95	0.96	0.93
Mean	0.94	0.93	0.94	0.90

Solar activity level is closely related to three types of sunspot groups, such as α, β, and $\beta\gamma$. From 2011 to 2016, there are a great number of α, β, and $\beta\gamma$ types. In particular, α and β types have hundreds of numbers monthly, where the sun is relatively active during this period. Since 2016, these types of sunspot groups have gradually decreased, even down to single digits, where the solar activities decrease. On the contrary, the numbers of $\beta\delta$ and $\beta\gamma\delta$ types have a little change, where the monthly amounts are only 12 and 18 at the peak, respectively.

The numbers of various types differ significantly in the same period of time, especially during the solar active period. For example, in November 2011 (the left-dashed vertical line), the numbers of α, β, $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ are 76, 144, 29, 5, and 7, respectively. The maximum difference (between β and $\beta\delta$) is up to 139. But, with the solar activity weakening, the difference becomes smaller. For example, in February 2020 (the right vertical line), the number of α, β, $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ are 1, 1, 0, 0, and 0, respectively. The maximum difference (between β and $\beta\delta$) is only 1.

5.4. The Relationship between Magnetic Classification and Solar Flares.

Solar flares are closely related to the magnetic classification of sunspot groups. The solar flare productivity is used to quantify the relationship, which is calculated by the total number of flares divided by the number of sunspot groups. According to the peak fluxes of soft X-ray, flares are usually classified into four levels: B, C, M, and X. Within a certain time interval, the total importance, I_{tot}, is presented as [48, 49]

$$I_{tot} = 0.1 \times \sum B + 1.0 \times \sum C + 10 \times \sum M + 100 \times \sum X.$$

(2)

The flare productivity with $I_{tot} \geq 1$, which is equivalent to a C1.0 flare, is analyzed. The occurrence of flares in an AR described in this paper is within 48 hours. The solar flare data come from the NOAA website (https://www.solarmonitor.org/). The number of sunspot groups and the corresponding flare productivity from January 2011 to April 2020 are listed in Table 3. They are also plotted in Figure 6 in order to make
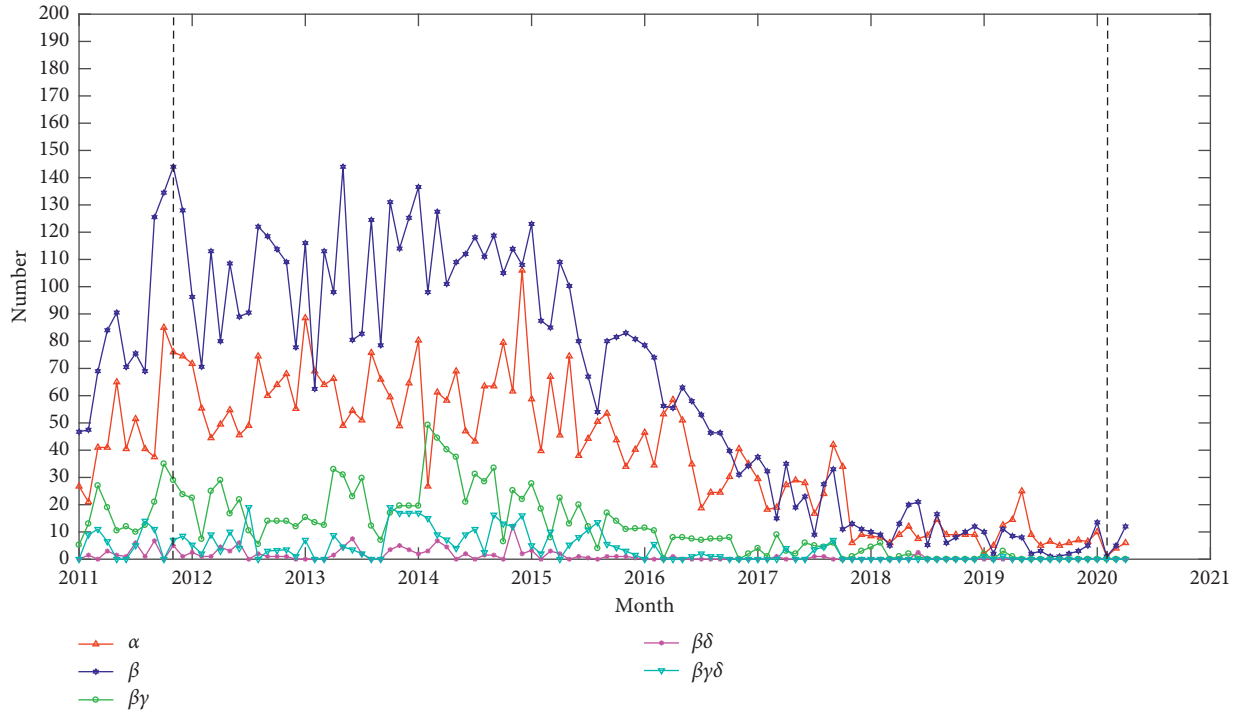
FIGURE 5: The total number for each type of sunspot groups in months from January 2011 to April 2020. The left dashed vertical line (November 2011) represents that the numbers of various types differ significantly, where the maximum difference is up to 139. The right dashed vertical line (February 2020) represents the numbers of various types that are very close, where the maximum difference is only 1.

TABLE 3: The number of various types of the sunspot groups and the corresponding flare productivity from 2011 to 2020 ($I_{tot} \geq 1$).

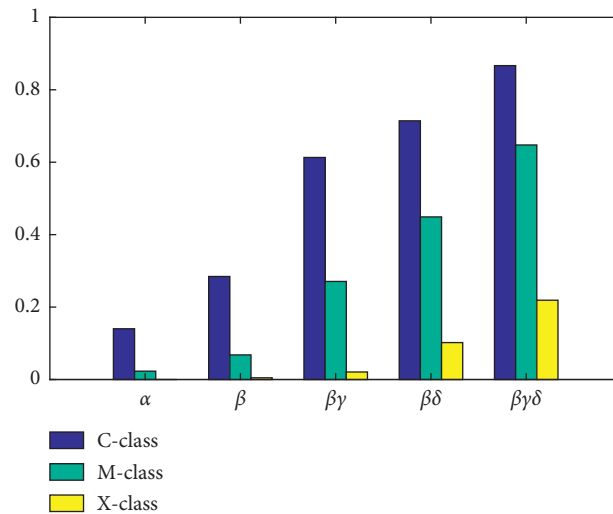| Type	Numbers of sunspot groups	Numbers of C-class flares	Numbers of M-class flares	Numbers of X-class flares	Productivity of C-class flares	Productivity of M-class flares	Productivity of X-class flares
α	649	91	15	0	0.14	0.02	0
β	1044	297	71	5	0.28	0.07	0.01
$\beta\gamma$	336	206	91	7	0.61	0.27	0.02
$\beta\delta$	49	35	22	5	0.71	0.45	0.10
$\beta\gamma\delta$	105	91	68	23	0.87	0.65	0.22
Total	2183	720	267	40	0.33	0.12	0.02

FIGURE 6: Productivity of C-class, M-class, and X-class flares of various types of the sunspot groups from January 2011 to April 2020 ($I_{tot} \geq 1$).

the relationship between magnetic classification of sunspot groups and flare productivities more intuitive.

It is found that the more complex the sunspot group is, the higher the corresponding flare productivity is. For instance, the flare productivities of C, M, and X belonging to the α type are all lower than those belonging to the β types (including β and $\beta\gamma$ types), and those of the β types are all lower than those of the δ types (including $\beta\delta$ and $\beta\gamma\delta$ types). In detail, the flare productivities of C-class, M-class, and X-class belonging to α are 0.14, 0.02, and 0, respectively; and those belonging to β are 0.28, 0.07, and 0.01, respectively. The flare productivities of these two types of sunspot groups are relatively low; in particular, X-class is almost zero. It means that these two types of sunspot groups with simple structures are difficult to produce large flares. On the other hand, the flare productivities of $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ are relatively high. The values for C-class flares reach up to 0.61, 0.71, and 0.87, respectively, and those for M-class flares are 0.27, 0.45, and 0.65, respectively. The X-class flares are very rare, as only forty occurred in the last ten years. The corresponding flare productivities of $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ are 0.02, 0.10, and 0.22, respectively. Especially, for a total of 105 $\beta\gamma\delta$, there erupted 23 X-class flares. That means $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ have much higher probabilities of flare eruption than the other types, especially $\beta\gamma\delta$. In additional, X-class flares basically occur above the δ or γ types.

The above results show that the Mount Wilson magnetic classification of sunspot groups is indeed closely related to solar flares. Therefore, the magnetic classification can be used as a major factor to predict the flare eruptions.

6. Discussion

6.1. Comparison with NOAA. We compared our results with NOAA carefully. Figure 7 shows the evolving procedure and classification results of NOAA 11158 from 13 to 15 February 2011.

It can be seen that the sunspot group during the three days is mixed in polarity, with opposite polarity within the same penumbra and no more than 2° separation of the umbra in the continuum image and the corresponding magnetogram. According to the description of Mount Wilson magnetic classification, the sunspot group should be $\beta\gamma\delta$. The results obtained by our method are all $\beta\gamma\delta$, while NOAA gives $\beta\gamma$ on the 13th and the 14th and $\beta\gamma\delta$ on the 15th. X-class large flares erupted above NOAA 11158 on the 14th. Our method classifies it into $\beta\gamma\delta$ two days earlier than NOAA and one day earlier than the eruption of this large flare.

Figure 8 shows the evolving procedure and classification results of NOAA 12260 from 8 to 10 January 2015. There is an AR visible at the lower right corner of NOAA 12258 on 8 January 2015. Our method detects the AR and classifies it into β in time. This AR is marked as NOAA 12260 and classified into β by NOAA on the 9th. Our method detects the AR one day earlier than NOAA.

Moreover, we counted the numbers of various types of sunspot groups detected by our method and NOAA in 2014 separately, as shown in Figure 9. The types of sunspot groups

in this year are very abundant. There are a few sunspots in some days and very big sunspots in other days. Importantly, a lot of different levels of flares erupted in that year. Therefore, it is a typical year that we chose for analyzing the magnetic classification of sunspot groups day by day. The x-axis represents the classification of sunspot groups, and the y-axis represents the number of each type. Our results are very similar to those of NOAA. The total numbers detected by our method and NOAA are 2254 and 2073, respectively. In detail, there are 654 sunspot groups belonging to α type by our method, 57 more than NOAA; 1155 β sunspot groups, 129 more than NOAA; 302 $\beta\gamma$ sunspot groups, 8 less than NOAA; 23 $\beta\delta$ sunspot groups, 4 more than NOAA; and 120 $\beta\gamma\delta$ sunspot groups, 1 less than NOAA. The main reason why there are more α and β types by our method than those by NOAA is that our method detects more of small α and β sunspot groups. On the other hand, the sum of the $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ sunspot groups detected by our method is 445, and that by NOAA is 450. The total difference is very small, and the difference of every type is also very small. We found that the difference is mainly caused by the somewhat similar features of these types during the evolution of sunspot groups (see Figure 7). Note that the statistical results are obtained by 313 days in 2014 due to missed data.

Figure 10 shows the flare productivities by our method and NOAA in 2014. Here, the total flare level, I_{tot}, is set to 10 or greater, which is equivalent to an M1.0 flare. The flares productivities of α, β, $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ detected by our method are 0.03, 0.10, 0.37, 0.60, and 0.64, respectively. They are 0.05, 0.15, 0.37, 0.40, and 0.48 by NOAA, respectively. The general trends of the flare productivities of different types by our method and NOAA are similar. The more complex the sunspot group is, the higher its flare productivity is. In other words, the flare productivities of $\beta\gamma\delta$, $\beta\delta$, and $\beta\gamma$ are both higher than that of β, and that of β is higher than that of α. If we have to say something different, NOAA obtains higher flare productivities of α and β compared to our method; our method has the same flare productivity as NOAA for $\beta\gamma$; for $\beta\delta$ and $\beta\gamma\delta$ that are most closely related to flares, the flare productivities of our method are higher than those of NOAA.

Note that the numbers of sunspot groups are counted somewhat differently in Figures 9 and 10. The number of sunspot groups in Figure 9 is the number of occurrences of sunspot groups in days, but in Figure 10, it is the number of different types of sunspot groups in their evolutions. For example, NOAA 11946 appeared on 5 January 2014 and disappeared on the 12th. It belongs to β in three days and to $\beta\gamma$ in five days. The numbers of β and $\beta\gamma$ in Figure 9 are 3 and 5, respectively, while both are 1 in Figure 10.

In summary, the multilevel features of various types of sunspot groups are extracted effectively using the deep learning method. This leads to a good performance in the magnetic classification of sunspot groups. Moreover, the complex sunspot groups including $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ have higher flare productivities, especially $\beta\gamma\delta$. Most X-class solar flares erupt above the complex sunspot groups. The Mount Wilson magnetic classification of the sunspot groups can indeed be used as an important predictor to predict the eruption of solar flares.
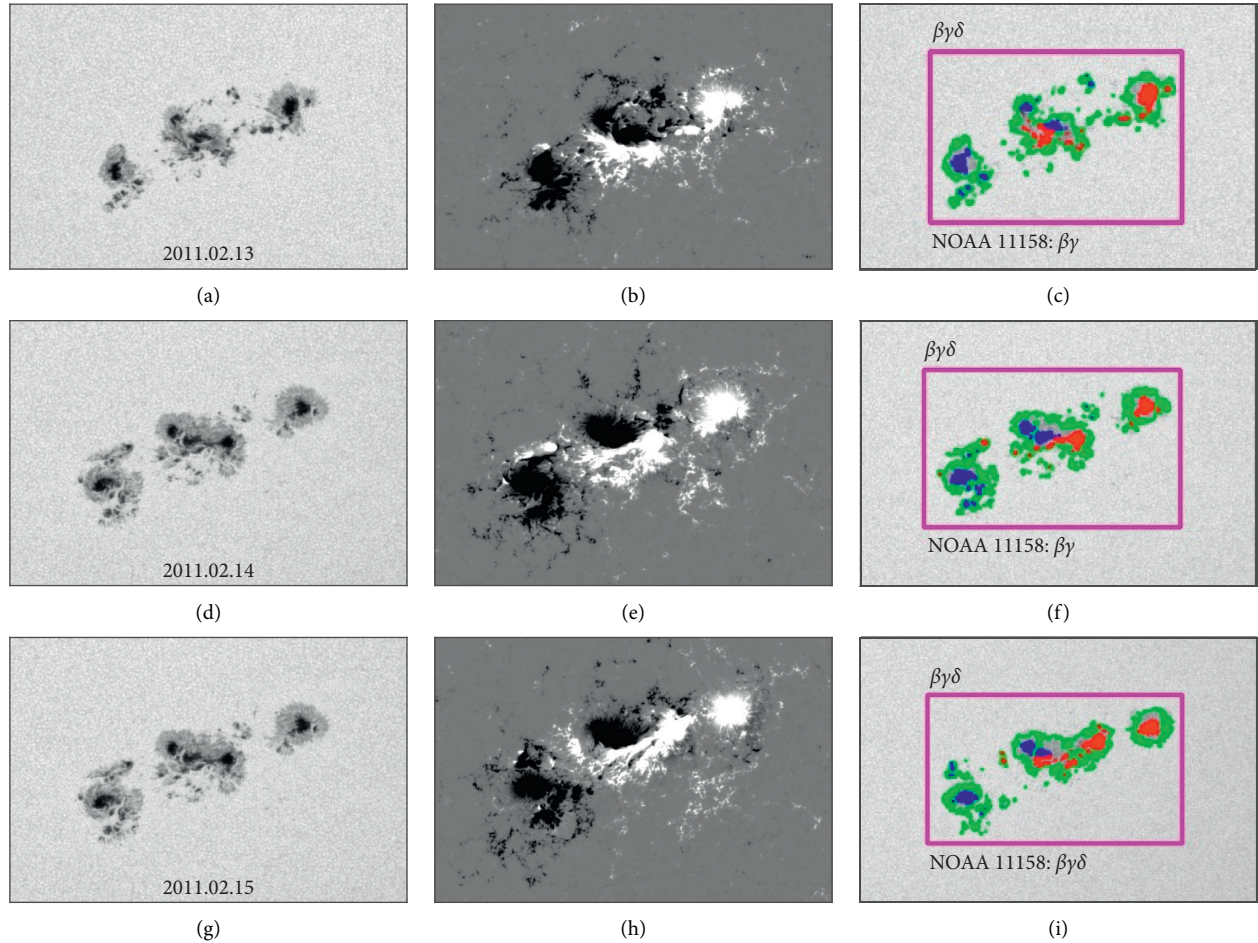
FIGURE 7: The classifications of NOAA 11158 given by our method and NOAA. The columns from left to right are the continuum image, magnetogram, and the classification result. The rows are the evolving procedure, which is on the 13th, 14th, and 15th of February 2011, respectively. All the times are 20:00:00 UT. The sunspot groups with solid-line boxes are detected by our method, where the classification results are displayed at the upper left of each box. The AR number and classification given by NOAA are manually labeled at the box below.

6.2. Some Problems.

Although our method has achieved a good performance in the Mount Wilson magnetic classification of sunspot groups, there are still some problems.

Some small spots are difficult to be affirmed because they are very similar to background. Figure 11(a) shows the result by our method on 2 August 2011. A total of four sunspot groups are detected and classified. Two regions with red circle labeled as 1 and 2 are zoomed in panels 11(b) and 11(c). Neither our method nor NOAA detects Region 1 as an AR, but both of them detect Region 2 as an AR (NOAA 11264). The difference is that our method classifies Region 2 into α, while NOAA does not give a classification result. It is not difficult to see that these two small spots are very similar. The main reason for miss-detected Region 1 is its environment. The reason is possibly due to the downsized images in order to reduce memory consumption. We believe this problem can be solved if the hardware conditions can support the high-resolution image without downsizing for deep learning operations.

Some sunspot groups are difficult to be judged when they are located at the edge of the solar surface. Figures 12(a) and 12(b) are subregions observed on May 10 2012 and 6 August

2012, respectively. In Figure 12(a), the sunspot group is about to disappear from the solar surface. On the other hand, the sunspot group in Figure 12(b) just appears on the solar surface. Both of them are classified into α type by our method. However, the characteristics of these sunspot groups are not fully shown. It is difficult to judge whether it is true. For such situations, it is not easy to classify them accurately even by hand.

Moreover, some sunspot groups are difficult to be classified during their evolution according to the definition of the Mount Wilson magnetic classification of sunspot groups. Figure 13 shows one common example, NOAA 12257, from 9 to 11 January 2015. The sunspot group is judged as $\beta\gamma$ by our method but as $\beta\delta$ by NOAA on 9 January 2015. We think that it should not be δ because of the umbra with the same polarity in the same penumbra. However, whether it belongs to β or $\beta\gamma$ is not easy to distinguish clearly. The main reason is that there are some unclear words in the definition, such as simple, complex, and irregular. For instance, β is a bipolar sunspot group, with the simple and obvious division between polarities; γ is a complex sunspot group with irregular polarity; $\beta\gamma$ is a bipolar sunspot group
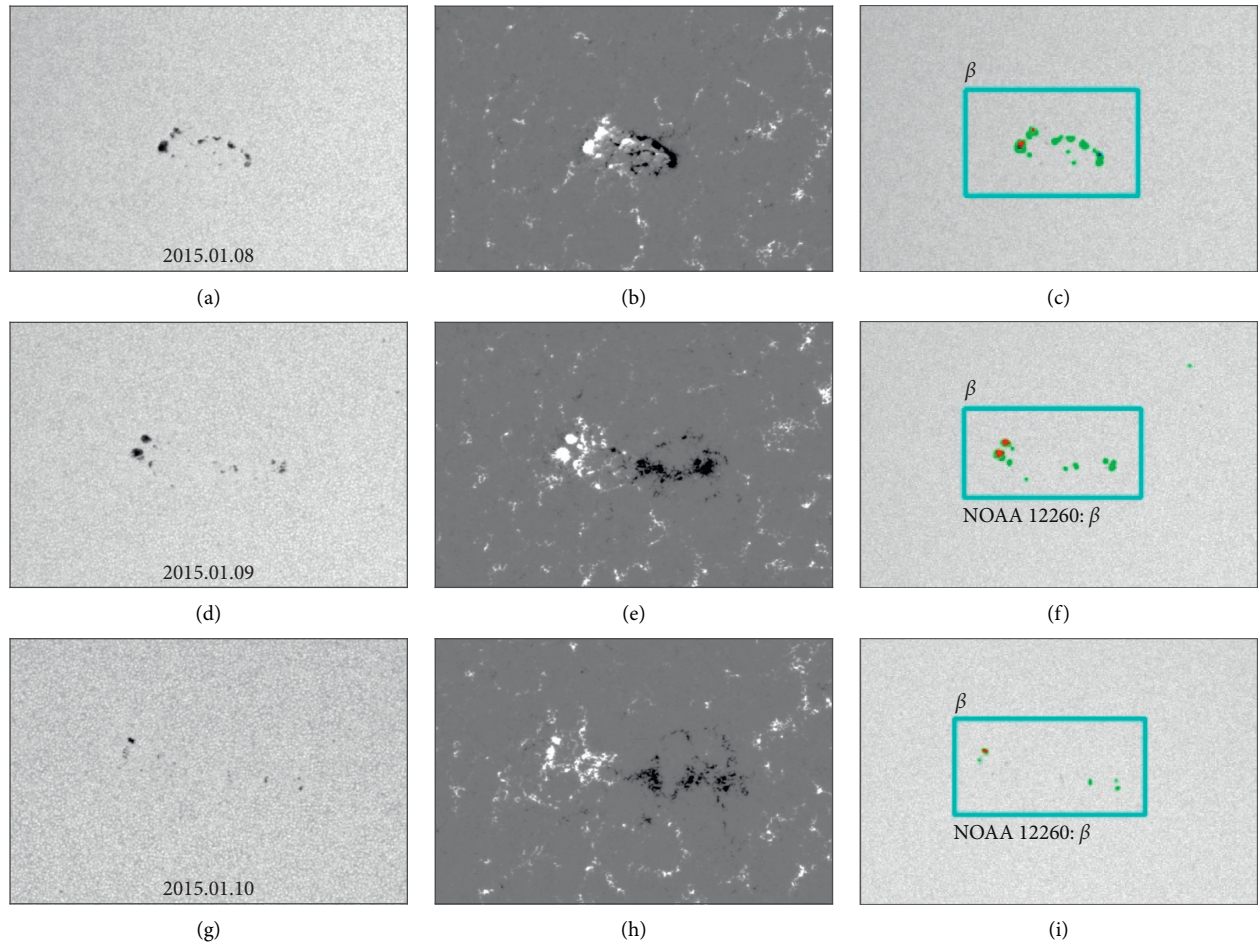
FIGURE 8: The classifications of NOAA 12660 given by our method and NOAA. The columns from left to right are the continuum image, magnetogram, and the classification result. The rows are the evolving procedure, which is on the 8th, 9th, and 10th of January 2015, respectively. All the times are 20:00:00 UT. The active region appeared on 8 January 2015. Our method detects the active region and classifies it into β in time. This AR is marked as NOAA 12260 and is classified into β on the 9th by NOAA.
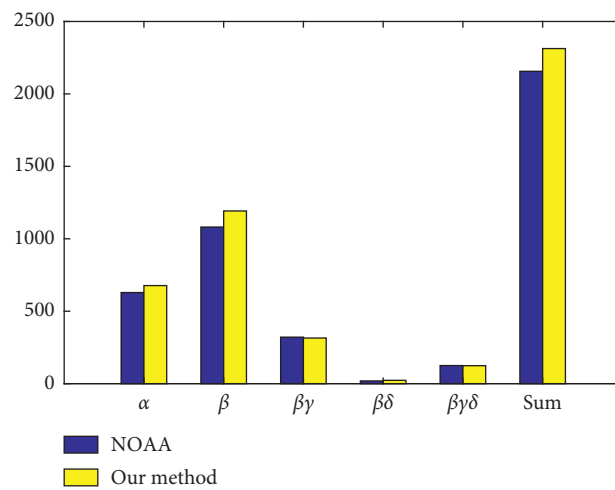
FIGURE 9: The numbers of various types of sunspot groups detected by our method and NOAA in 2014.

but with more than one continuous polarity reversal line. What level is simple or complex especially in the transitional period from one type into another type during its evolution?

Here, it is judged by our method to be $\beta\gamma$. Now, in the second row that follows, it is judged as δ because of the umbra with opposite polarity appear in the same penumbra, which are
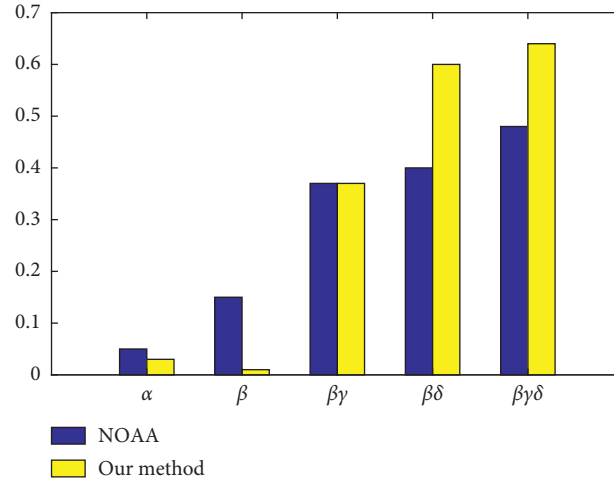
FIGURE 10: The flare productivities of various types of sunspot groups detected by our method and NOAA in 2014 ($I_{tot} \geq 10$).
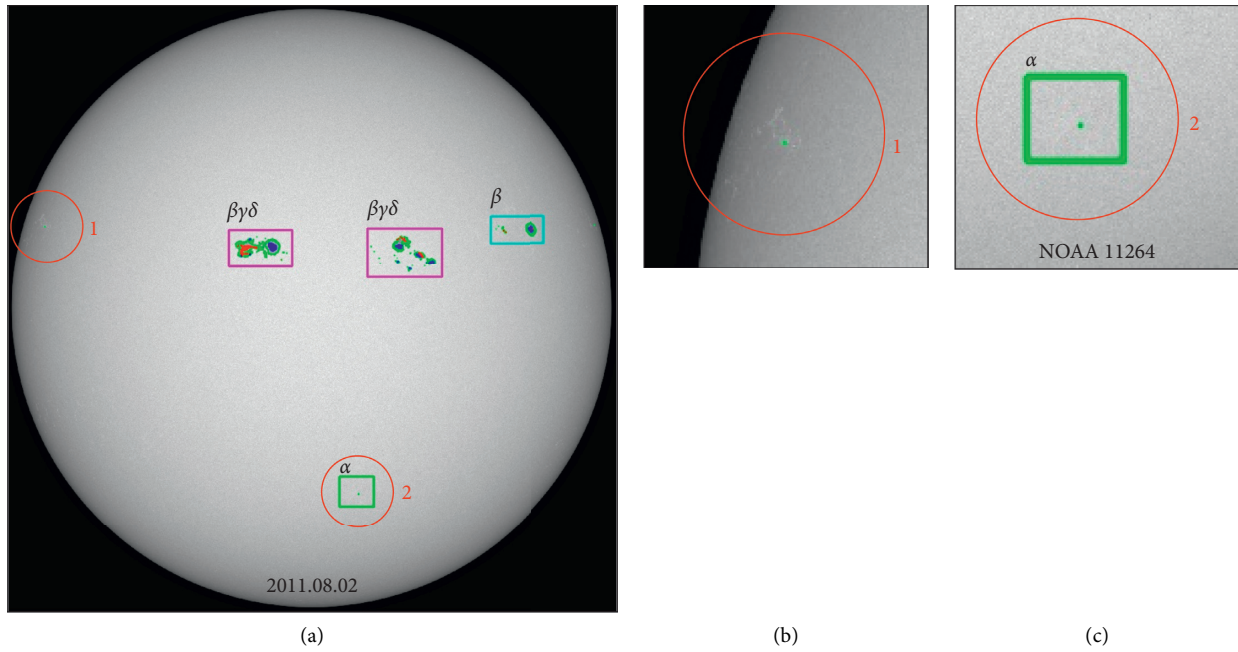

FIGURE 11: Detection results by our method at 20:00:00 UT on 2 August 2011. (a) The detection results of the composited solar full image. The region with a red circle labeled as 1 is not detected, which is zoomed in (b). The region labeled as 2 is detected, which is zoomed in (c).
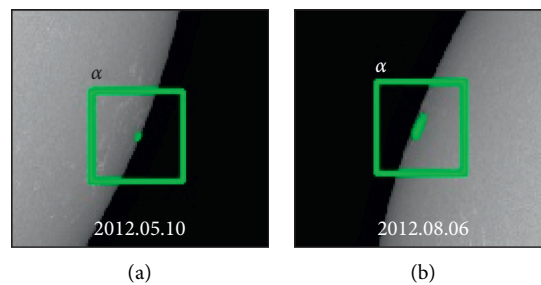

FIGURE 12: (a, b) Subregions at the edges of the solar surface observed at 20:00:00 UT on 10 May 2012 and at 20:00:00 UT on 6 August 2012, respectively.
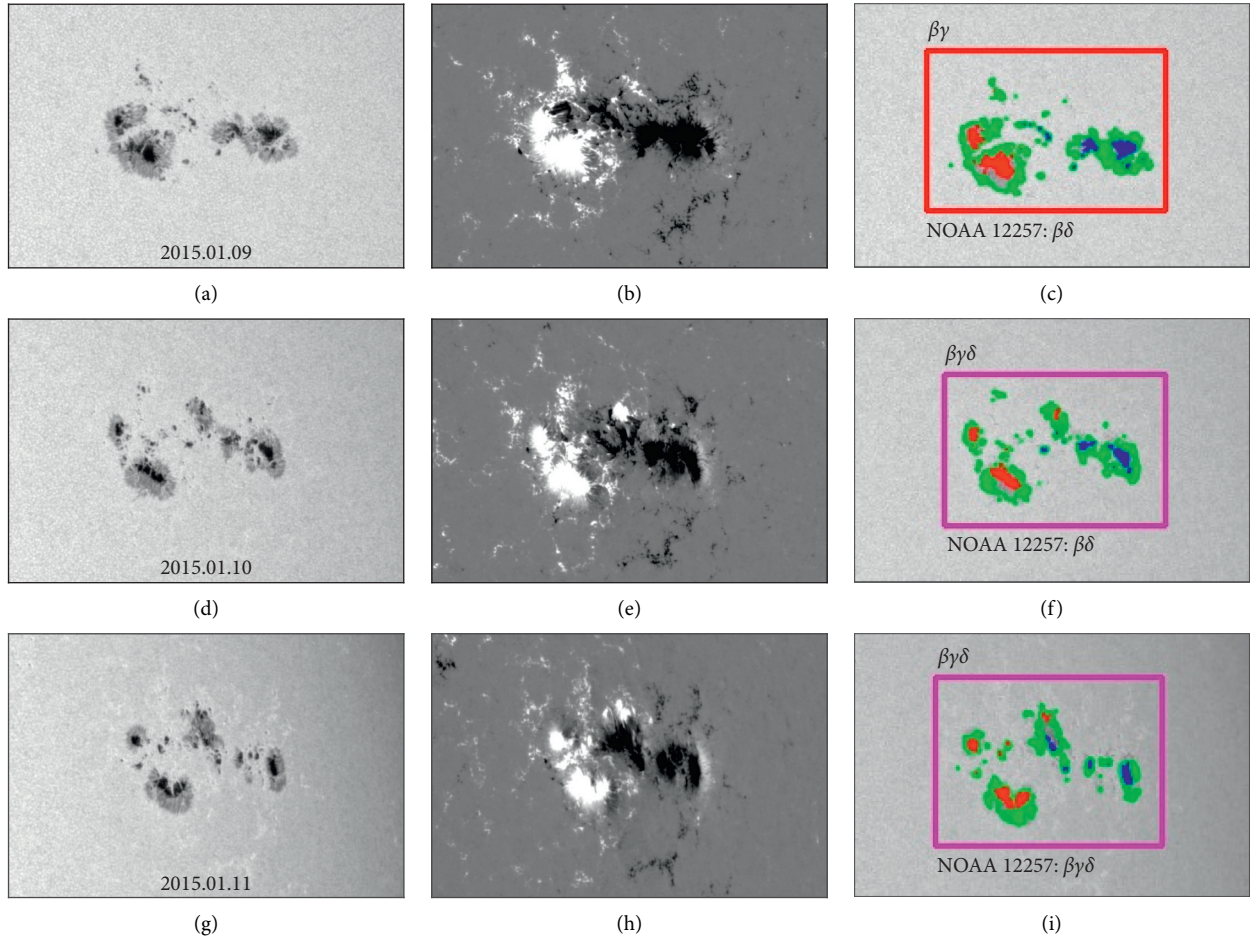
FIGURE 13: The classifications of NOAA 12257 given by our method and NOAA. The columns from left to right are the continuum image, magnetogram, and the detection result, respectively. The rows show the evolving procedure, which is on the 9th, 10th, and 11th of January 2015, respectively. All the times are 20:00:00 UT. The sunspot groups with solid-line boxes are detected by our method, where the classification results are displayed at the upper left of each box. The AR numbers and classification given by NOAA are manually labeled at the box below.

coincident by both methods. Whether it belongs to β or $\beta\gamma$ is also a problem. We think that the sunspot group is inclined to be $\beta\gamma\delta$ for its complex sunspot group with irregular polarity here. In the third row, both methods obtain the same results, $\beta\gamma\delta$.

7. Conclusion

This paper proposes a deep learning method for the Mount Wilson magnetic classification of sunspot groups. The continuum images and the magnetograms from HMI spanning from June 2010 to April 2020 and MDI spanning from July 2000 to September 2001 are employed to build the dataset. We first composite the morphological structures of sunspot groups from the continuum image and the magnetic properties from the cotemporal magnetogram into one image. A total of 2486 composited images are then labeled to build the dataset of the Mount Wilson magnetic

classification, which are divided into a training set and a test set at a ratio of 3 : 1.

A deep learning model, CornerNet-Saccade, is adopted to detect and classify the sunspot groups. It mainly includes three stages, generating object locations, detecting objects, and merging detections. The first stage is to find approximate locations and rough sizes of the sunspot groups on the solar full images. The backbone of the method is Hourglass-54, which consists of 3 hourglass modules with a total depth of 54 layers. Each hourglass module is a symmetrical structure, which downsamples to a very low resolution and then upsamples and combines features across multiple resolutions. The multiple hourglass modules together allow for repeated bottom-up, top-down inference across scales, which capture and consolidate information across all scales of the image. A total of three size scales of attention maps are predicted using the feature maps gotten from Hourglass-54 by the attention mechanism, corresponding to small,

medium, and large objects. Meanwhile, the bounding boxes are generated by detecting the two key points, the top-left corners and the bottom-right corners of the object, through corner pooling layers. The second stage is to further determine the specific locations and bounding boxes of the sunspot groups. The locations gotten from the first stage are ranked by their scores, and then the top ones are picked up. They are fine-adjusted by corner pooling layers using the feature maps through the second Hourglass-54. The third stage is to merge the detections by NMS for eliminating the redundant boxes. These key technologies improve the efficiency of the deep learning method without sacrificing accuracy.

The model is trained on a personal computer equipped with a GeForce RTX$^{\text{TM}}$ 2080 graphics card. After 335,000 iterations, the loss value converges stably. It takes about 3 seconds to detect and classify the sunspot groups in a composited solar full image. The experimental results show that this method has an excellent performance in the detection and classification. The accuracy, precision, recall, and mAP reach up to 0.94, 0.93, 0.94, and 0.90, respectively.

Based on the reliable performance of our method, the classifications of sunspot groups in the past 10 years from January 2011 to April 2020 are analyzed, which are shared on the website (https://61.166.157.71/MWMCSG.html). In addition, we calculate the flare productivities of various types of sunspot groups from January 2011 to April 2020 in order to describe the relationship between the magnetic classification of sunspot groups and solar flares. If I_{tot} is set to greater than or equal to 1, the flare productivities of α, β, $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ sunspot groups are 0.14, 0.28, 0.61, 0.71, and 0.87, respectively. If I_{tot} is set to greater than or equal to 10, they are 0.02, 0.07, 0.27, 0.45, and 0.65, respectively. Among them, the flare productivities of $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ reach up to 0.60 with $I_{\text{tot}} \geq 1$. The productivities of $\beta\delta$ and $\beta\gamma\delta$ reach up to 0.45 even with $I_{\text{tot}} \geq 10$. This means that $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ sunspot groups are indeed very closely related to solar flare eruptions, especially the $\beta\gamma\delta$ sunspot groups.

We also compare the classifications of sunspot groups with NOAA using the data in 2014. The results show that our method detects more α and β sunspot groups which are small and faint; therefore, the numbers of α and β are 654 and 1155, respectively, 57 and 129 more than those of NOAA, respectively. Our method is good at distinguishing the more complex sunspot groups including $\beta\gamma$, $\beta\delta$, and $\beta\gamma\delta$ classes, owing to the features of different types learned by the deep learning networks, where the detection schemes of CornerNet-Saccade, such as Hourglass-54, key points, and the attention map mechanism, make the features of the multiple levels fused available.

In the future, we plan to use the Mount Wilson magnetic classification of sunspot groups as the main predictor to establish a more reasonable solar flare prediction model.

Data Availability

The magnetogram and continuum image data used to support the findings of this study are observed by SDO/HMI and SOHO/MDI; all of the FITS files we have used are downloaded from https://jsoc.stanford.edu/. The SRS text files can be downloaded from the NOAA/SWPC website https://www.swpc.noaa.gov/. The training set and test set data for deep 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

References

[1] K. Sakurai, "Motion of sunspot magnetic fields and its relation to solar flares," *Solar Physics*, vol. 47, no. 1, pp. 261–266, 1976.

[2] P. O. Taylor, "The relationship between sunspot and solar flare activities for the period 1974–1987," *Journal of the American Association of Variable Star Observers*, vol. 17, pp. 20-21, 1988.

[3] Z. Shi and J. Wang, "Delta-sunspots and X-class flares," *Solar Physics*, vol. 149, no. 1, pp. 105–118, 1994.

[4] 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, pp. 1135–1140, 2005.

[5] 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.

[6] Y. Zhang, J. Liu, and H. Zhang, "Relationship between rotating sunspots and flares," *Solar Physics*, vol. 247, no. 1, pp. 39–52, 2008.

[7] A. G. Emslie, B. R. Dennis, A. Y. Shih et al., "Global energetics of thirty-eight large solar eruptive events," *The Astrophysical Journal*, vol. 759, no. 1, p. 71, 2012.

[8] M. Zhao, J. Chen, and Y. Liu, "Statistical analysis of sunspot groups and flares for solar maximum and minimum," *Scientia Sinica Physica, Mechanica & Astronomica*, vol. 44, pp. 109–120, 2014.

[9] K. O. Kiepenheuer, "Solar activity," in *The Sun*, p. 322, 1953.

[10] A. L. Cortie, "On the types of sun-spot disturbances," *The Astrophysical Journal*, vol. 13, p. 260, 1901.

[11] P. Simon, G. Heckman, and M. Shea, *COSPAR: Solar-terrestrial Predictions: Proceedings of a Workshop at Meudon, France, June 18-22, 1984*, National Oceanic and Atmospheric Administration, Washington, DC, USA, 1986.

[12] P. S. McIntosh, "The classification of sunspot groups," *Solar Physics*, vol. 125, no. 2, pp. 251–267, 1990.

[13] G. E. Hale, F. Ellerman, S. B. Nicholson, and A. H. Joy, "The magnetic polarity of sun-spots," *The Astrophysical Journal*, vol. 49, p. 153, 1919.

[14] X. L. Yan, L. H. Deng, Z. Q. Qu, and C. L. Xu, "The phase relation between sunspot numbers and soft X-ray flares," *Astrophysics and Space Science*, vol. 333, no. 1, pp. 11–16, 2011.

[15] G. R. Greatrix, "On the statistical relations between flare intensity and sunspots," *Monthly Notices of the Royal Astronomical Society*, vol. 126, no. 2, pp. 123–133, 1963.

[16] 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.

[17] S. Eren, A. Kilcik, T. Atay et al., "Flare-production potential associated with different sunspot groups," *Monthly Notices of the Royal Astronomical Society*, vol. 465, no. 1, pp. 68–75, 2017.

[18] T. T. Nguyen, C. P. Willis, D. J. Paddon, S. H. Nguyen, and H. S. Nguyen, "Learning sunspot classification," *Fundamenta Informaticae*, vol. 72, pp. 295–309, 2006.

[19] M. A. Abd, S. F. Majed, and V. Zharkova, "Automated classification of sunspot groups with support vector machines," in *Technological Developments in Networking, Education and Automation*, K. Elleithy, T. Sobh, M. Iskander, V. Kapila, M. A. Karim, and A. Mahmood, Eds., Springer, Dordrecht, Netherlands, 2010.

[20] T. Colak and R. Qahwaji, "Automated Solar Activity Prediction: a hybrid computer platform using machine learning and solar imaging for automated prediction of solar flares," *Space Weather*, vol. 7, no. 6, 2009.

[21] S. Hong, J. Kim, J. Han, and Y. Kim, "The automatic solar synoptic analyzer and solar wind prediction," *AGU Fall Meeting Abstracts*, vol. 2014, pp. SH21A–4089, 2014.

[22] S. Padinhatteeri, P. A. Higgins, D. Shaun Bloomfield, and P. T. Gallagher, "Automatic detection of magnetic δ in sunspot groups," *Solar Physics*, vol. 291, no. 1, pp. 41–53, 2016.

[23] Y. Fang, Y. Cui, and X. Ao, "Deep learning for automatic recognition of magnetic type in sunspot groups," *Advances in Astronomy*, vol. 2019, pp. 1–10, 2019.

[24] 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. 2278–2324, 1998.

[25] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[26] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[27] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," 2013, https://arxiv.org/abs/1311.2524.

[28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," 2015, https://arxiv.org/abs/1506.02640.

[29] H. Law, Y. Teng, O. Russakovsky, and J. Deng, "CornerNet-Lite: efficient keypoint based object detection," 2019, https://arxiv.org/abs/1904.08900.

[30] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, https://arxiv.org/abs/1409.0473.

[31] L. Tychsen-Smith and L. Petersson, "DeNet: scalable real-time object detection with directed sparse sampling," 2017, https://arxiv.org/abs/1703.10295.

[32] X. Wang, K. Chen, Z. Huang, C. Yao, and W. Liu, "Point linking network for object detection," 2017, https://arxiv.org/abs/1706.03646.

[33] H. Law and J. Deng, "CornerNet: detecting objects as paired keypoints," 2018, https://arxiv.org/abs/1808.01244.

[34] 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.

[35] P. H. Scherrer, R. S. Bogart, R. I. Bush et al., "The solar oscillations investigation-Michelson Doppler imager," *Solar Physics*, vol. 162, no. 1-2, pp. 129–188, 1995.

[36] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman, "LabelMe: a database and web-based tool for image annotation," *International Journal of Computer Vision*, vol. 77, no. 1–3, pp. 157–173, 2008.

[37] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," 2016, https://arxiv.org/abs/1603.06937.

[38] O. Rukundo and H. Cao, "Nearest neighbor value interpolation," 2012.

[39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, https://arxiv.org/abs/1512.03385.

[40] B. W. Tatler, N. J. Wade, H. Kwan, J. M. Findlay, and B. M. Velichkovsky, "Yarbus, eye movements, and vision," *I-Perception*, vol. 1, no. 1, pp. 7–27, 2010.

[41] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "generalized intersection over union: a metric and a loss for bounding box regression," 2019, https://arxiv.org/abs/1902.09630.

[42] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," 2017, https://arxiv.org/abs/1708.02002.

[43] R. Girshick, "Fast R-CNN," 2015, https://arxiv.org/abs/1504.08083.

[44] A. Newell and J. Deng, "Pixels to graphs by associative embedding," 2017, https://arxiv.org/abs/1706.07365.

[45] A. Neubeck and L. Gool, "Efficient non-maximum suppression," in *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*, pp. 850–855, Hong Kong, August 2006.

[46] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," 2014, https://arxiv.org/abs/1412.6980.

[47] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (VOC) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.

[48] A. Antalova, "Daily soft X-ray flare index (1969–1972)," *Contributions of the Astronomical Observatory Skalnate Pleso*, vol. 26, pp. 98–120, 1996.

[49] V. I. Abramenko, "Relationship between magnetic power spectrum and flare productivity in solar active regions," *The Astrophysical Journal*, vol. 629, no. 2, pp. 1141–1149, 2005.