RESEARCH

Solar Active Regions Detection and Tracking Based on Deep Learning

Long Gong[1] · Yunfei Yang[1] · Song Feng[1] · Wei Dai[1] · Bo Liang[1] · Jianping Xiong[2]

Abstract
Solar active regions serve as the primary energy sources of various solar activities, directly impacting the terrestrial environment. Therefore precise detection and tracking of active regions are crucial for space weather monitoring and forecasting. In this study, a total of 4577 HMI and MDI longitudinal magnetograms are selected for building the dataset, including the training set, validating set, and ten testing sets. They represent different observation instruments, different numbers of activity regions, and different time intervals. A new deep learning method, ReDetGraphTracker, is proposed for detecting and tracking the active regions in full-disk magnetograms. The cooperative modules, especially the redetection module, NSA Kalman filter, and the splitter module, better solve the problems of missing detection, discontinuous trajectory, drifting tracking bounding box, and ID change. The evaluation metrics *IDF1*, *MOTA*, *MOTP*, *IDs,* and *FPS* for the testing sets with 24-h interval on average are 74.0%, 74.7%, 0.130, 13.6, and 13.6, respectively. With the decreasing intervals, the metrics become better and better. The experimental results show that ReDet-GraphTracker has a good performance in detecting and tracking active regions, especially capturing an active region as early as possible and terminating tracking in near-real time. It can well deal with the active regions whatever evolve drastically or with weak magnetic field strengths, in a near-real-time mode.

1. Introduction

Solar active regions (ARs) are the regions on the Sun with strong magnetic fields. These regions are where most of solar activities, like solar flares and solar eruption as coronal mass ejections (CME), come from. Consequences of such solar activities extend to terrestrial

---

✉ Y. Yang
yangyf@kust.edu.cn

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 Yunnan Observatories, Chinese Academy of Sciences, Kunming 650051, China

space environments, manifesting as interference in electromagnetic communication channels, degradation of electrical power infrastructure, or perturbations in radio signal transmission. Therefore detection and tracking solar active regions in near-real time is meaningful for human space exploration and life on Earth (Qahwaji and Colak 2005; Zhang, Wang, and Liu 2010).

In the past years, the detection and tracking methods for solar active regions mainly used traditional image processing technologies (Gallagher, Moon, and Wang 2002; LaBonte, Georgoulis, and Rust 2007; Colak and Qahwaji 2009). The detection method usually depended on thresholds (McAteer et al. 2005; Zhang, Wang, and Liu 2010; Caballero and Aranda 2014), such as intensity threshold, opening and closing operator threshold, or boundary threshold of region growth. The thresholds generally were set as parameters by experience or multiple experiments, which are crucial for these methods. Some studies discussed the parameters in detail and then gave some guidelines for setting parameters (Zharkova et al. 2005; Zhang, Wang, and Liu 2010; Higgins et al. 2011; Caballero and Aranda 2014). Especially, the National Oceanic and Atmospheric Administration (NOAA) numbers are widely approved around the world (Gallagher, Moon, and Wang 2002). The NOAA active regions are identified manually using three overlays: a Stoneyhurst overlay, a sunspot-area overlay, and a limb-area correction overlay (Giersch, Kennewell, and Lynch 2018; Meadows 2020). Also widely known, the Space-Weather HMI Active Region Patches (SHARPs: Turmon et al. 2011; Hoeksema et al. 2014; Bobra et al. 2021) and the Space-Weather MDI Active Region Patches (SMARPs: Bobra et al. 2021) were produced by similar pipeline described in (Turmon, Pap, and Mukhtar 2002; Turmon et al. 2010). They computed a pixel-scale activity mask by a Bayesian approach first and then grouped it into NOAA AR-scale regions by a matched filter approach.

Tracking solar active regions mainly predicted the positions at the next frame based on the differential rotation theorem (Thompson et al. 2003) and then performed target association based on Euclidean distance, feature correlation, or overlap score (Turmon, Pap, and Mukhtar 2002; Turmon et al. 2010; Kempton, Pillai, and Angryk 2014; Bobra et al. 2021). In 2023 an automatic tracking algorithm (AutoTAB) was proposed for the solar bipolar magnetic regions (Sreedevi et al. 2023). They detected solar bipolar magnetic regions based on threshold and magnetic flux balance conditions and then tracked them, still mainly based on differential rotation theorem. The HARPs (the Geometry of HMI Active Region Patches) have two kinds of products, the definitive data and the near-real-time (NRT) data. The definitive data products will be provided after a couple days (generally a half month) from the observed time until the definitive observables are complete. Since the entire history of the HARP is known, each definitive HARP encloses the same heliographic area during its entire lifetime while keeping relatively coherent evolution. For quick-look purposes, the NRT HARPs are provided after several hours of observation. The NRT HARP will be labeled once it is identified. Its size will change depending on its real status. Also, it may merge, resulting in termination of one or more NRT HARPs and continuation of the larger merged object. Note that all the definitive SMARPs products from 1996 to 2010 were provided by Bobra et al. (2021) and the NRT data do no need to be provided.

In recent years, with the popularization of deep learning, some deep learning detection and tracking algorithms have also been applied for some targets on the Sun, such as sunspot identification (Yang et al. 2018; Mourato, Faria, and Ventura 2024), filament detection (Guo et al. 2022; Zheng et al. 2024), bright point detection (Yang et al. 2019; Xu et al. 2021; Bai et al. 2023), Hα fibrils tracking (Jiang et al. 2021), and solar flares tracking (Oludehinwa et al. 2023).

Therefore we propose a new deep learning method, ReDetGraphTracker, as an end-to-end system, which combines the detection and tracking active regions together. Using the

Table 1 The information of all samples with different intervals.

Interval	Training set		Testing set		Validating set		Total
	number	percentage	number	percentage	number	percentage	
24 h	13,943	60%	2896	60%	2896	60%	19,306
6 h	4648	20%	965	20%	965	20%	6435
3 h	2324	10%	483	10%	483	10%	3218
12 m	2324	10%	482	10%	482	10%	3218
Total	23,239	–	4826	–	4826	–	32,891

full-disk longitudinal magnetograms, active regions will be detected, tracked, and terminated in time only depending on its features in image and evolution in image series by deep-learning model. The active regions will keep coherent evolution as far as possible in near-real-time mode, which do not need to know their entire history. For improving the accuracy of the detection, a redetection module was designed to supplement the base detection module. These two detection modules respectively focus on active regions with obvious magnetic characteristics and relatively weak ones. For tracking active regions with high quality as far as possible, a splitter module was used for reducing the trajectory interruption, and then a flow graph in the AR association module was used for associating active regions between successive frames.

This paper is structured as follows. Section 2 introduces the data source. Section 3 details the proposed detection and tracking deep learning methods. Section 4 shows the configuration and parameters of experiments. Section 5 analyzes and discusses the results. Section 6 briefly summarizes the work.

2. Data Set

This work employs the full-disk 720-second line-of-sight (LoS) magnetic field maps (hmi.M_720s) from the Helioseismic and Magnetic Imager (HMI: Scherrer et al. 2012) on the Solar Dynamics Observatory (SDO) and the full disk 96-minute LoS magnetograms (mdi.fd_M_96m_lev182) from the Michelson Doppler Image (MDI: Turmon, Pap, and Mukhtar 2002) on the Solar and Heliospheric Observatory (SOHO). The training data are all from HMI. The MDI data are adopted into the test set for verifying the generalization of the method. A total of 4577 full-disk solar longitudinal magnetograms from 12 July 2000 to 15 May 2024 are representatively selected. The data of training set are selected from 1 January 2015 to 30 April 2018. The test set data are selected from nonoverlapping time periods to ensure no overlap with the training set. The time intervals are set for adapting to different interval tracking requirements. Table 1 lists the information of all data sets with different intervals. A total of 32,891 samples are labeled, which are divided into training set, validating set, and testing set in 70: 17: 13. The time intervals are set as 24 hours, 6 hours, 1 hour, and 12 minutes, with their proportions being 60%, 20%, 10%, and 10%, respectively. The video sequences have unfixed frame number; the maximum frame number only relies on the memory of GPU. These samples are annotated by DarkLabel (darkpgmr 2020), an open-source video-annotation software, which is used to make annotation data by annotating the location and identity ID of each target across video sequences. Note that the sequences need video format, such as MPEG, AVI, MP4. Here we converted these sets of image sequences to AVI format. Finally, the annotation data are converted to a mature

Table 2 Testing set.

Dataset	Data source	Time	Interval	Frame
D_1	MDI	2010.07.01 – 2010.07.31	24 h	30
D_2	MDI	2000.07.12 – 2000.08.13	24 h	30
D_3	HMI	2010.07.01 – 2010.07.31	24 h	30
D_4	HMI	2010.11.05 – 2010.12.04	24 h	30
D_5	HMI	2024.03.01 – 2024.05.15	24 h	45
D_6	HMI	2014.10.01 – 2014.10.31	24 h	30
D_{6-1}	HMI	2014.10.01 – 2014.10.31	6 h	120
D_{6-2}	HMI	2014.10.01 – 2014.10.31	1 h	720
D_{6-3}	HMI	2014.10.01 – 2014.10.31	12 m	1440
D_{6-4}	HMI	2014.10.01 – 2014.10.31	48 h	15

labeling format: Multi-Object Tracking benchmark (Milan et al. 2016). Determining an active region is mainly according to SolarMonitor (www.SolarMonitor.org), supplemented by manual assistance.

The testing set consists of ten sets of image sequences, which respectively represent different observation instruments, different numbers of activity regions, and different time intervals. Table 2 lists the information of the testing sets. Both D_1 and D_3 include fewer active regions, whereas both D_2 and D_5 include a larger number of active regions. D_1 from MDI and D_3 from HMI are selected with the same time period. D_5 is the newest data that could download today. Considering the tracking requirement in day or less, the interval of sets from D_1 to D_5 is set as 24 h. With evolving, the active regions change usually various after a day. Within a day, the smaller interval of image sequences will make the tracking easier and more accurate in theory. Therefore we adopted a same image sequence that corresponds a very active Sun to build three test sets from D_{6-1} to D_{6-3} with intervals as 6 h, 1 h and 12 m, respectively. Additionally, a set with 48 h, D_{6-4}, is also implemented. Note that the time of images in the sets with intervals as 24 h or 48 h are basically consistent with that provided by SolarMonitor for comparison.

3. Method

A new deep-learning method was designed for detection and tracking active regions, named as ReDetGraphTracker, which was built as an end-to-end system. It is mainly composed of two parts, detection and tracking modules. The whole structure is shown in Figure 1. The detection module focuses on extracting features, which usually include shape, texture, and other information related to the active regions. The tracking module implements data association between active regions in successive frames for continuous tracking of the same active region. The main tasks are matching the appearance features and motion information of the active regions, continuous updating the state of active regions and trajectory prediction based on historical movement information for providing accurate data, and enhancing the understanding of behavior, including position, speed, size, and other details.
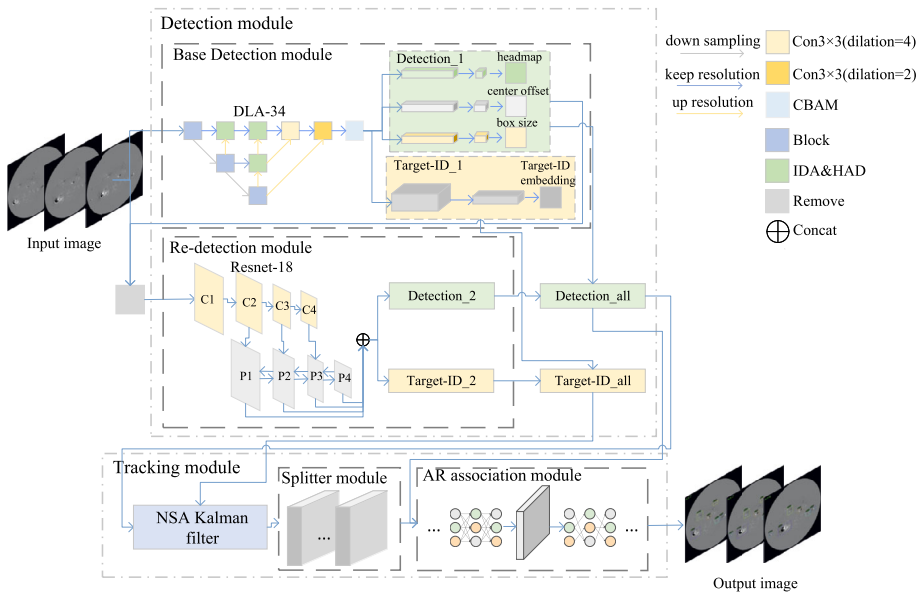
Figure 1 The structure of ReDetGraphTracker method.

3.1. Detection Module

3.1.1. Base Detection Module

The detection module consists of base detection module and redetection module. The base detection module uses a sequence of magnetograms as input, extracts the features through a backbone network, and then outputs to the detection branch and the target-ID information branch, respectively.

Some outstanding backbones have been proposed in recent year, such as ResNet (He et al. 2016), YOLO (Redmon et al. 2016), and so on. Considering the balance between speed and accuracy, we adopted DLA-34 (Yu et al. 2018), an encoder-decoder structure, as the backbone for base detection. It firstly uses three blocks to extract feature maps with different sizes and channels. Each block in a sequence includes a 3×3 convolution, a batch normalization layer, a ReLU activation function, and a 3×3 convolutional batch normalization layer. Then both feature maps through block i and block $i - 1$ are subjected to a module integrating iterative deep aggregation and hierarchical deep aggregation (HDA and IDA, Yu et al. 2018). IDA fuses shallow and deep semantic information, and HDA integrates channel information. Finally, after 3×3 convolution with hole rates of 2 and 4, the aggregated feature maps are transformed to a convolutional attention mechanism module (CBAM: Woo et al. 2018) to obtain more global information.

The detection branch consists of three prediction heads, namely heatmap head, center offset head, and box size head. The structures of these three prediction heads are the same, including a $256 \times 3 \times 3$ convolution and a $2 \times 1 \times 1$ convolution. The heatmap prediction head is to obtain the corresponding heatmap information for predicting accurate target information. Center offset head is to obtain the target position information for predicting the center point coordinates of the active regions. The box size head obtains detection box

information. They are packaged as detection information, after filtering by the nonmaximum suppression algorithm (Gallagher, Moon, and Wang 2002). The target-ID information branch obtains target-ID embedding through two convolutions, which is a one-dimensional vector representing the initial state. It will be used for the ID matching during tracking.

3.1.2. Redetection Module

No matter how excellent the detection module is, it is not possible to completely avoid wrong or missing detections. The wrong detection still has a chance to be discarded during tracking; however, the missing detection will cause discontinuous tracking target which is difficult to solve during tracking. Therefore we designed a redetection module for remedying those missing detections. The lightweight Resnet-18 (He et al. 2016) was selected as a backbone and fused by the Path aggregation Feature Pyramid Network (PAFPN: Liu et al. 2018).

The key of the redetection module is the input images, in which all the detections outputted from the base detection module have been removed. In detail, those pixels belonging to the targets detected in the procedure of the base detection module will be set as zero. The purpose of doing this is to make the images no longer include these detected targets. It is feasible that those undistinguished features could be captured in the redetection module. These images are sent to Resnet-18 for feature extraction with different scales, and then PAFPN selects feature maps C4, C3, and C2 from top to bottom for feature fusion. Comparing with classical FPN (Lin et al. 2017a), PAFPN adds a bottom-up path after the top-down path and replaces shortcut connections with concatenation. The structure of the prediction heads is the same as that of the base detection module. The concatenated feature maps are transferred to three prediction heads and target-ID branch. The prediction results are also packaged as detection information after filtering by NMS algorithm.

The detection information and target-ID embedding obtained by the base detection module and those by the redetection module are integrated, which can remedy the missing detections during the base detection procedure as far as possible. The final detection information and final target-ID information are generated, which are called detection_all and target-ID_all, respectively.

3.2. Tracking Module

The tracking module includes three parts: a noise scale adaptive (NSA) Kalman filter, a splitter module, and an AR association module.

3.2.1. NSA Kalman Filter

The detection_all and target-ID_all are inputted into the NSA Kalman filter (Du et al. 2021), which will predict the position information and motion information of each tracked active region. The position information conveys central point coordinates of an active region, specifying its exact position in space. The motion information includes target-ID_all, the velocity information, the trajectory information, and the time stamp. Among them, trajectory information is a series of two-dimensional data that describes the change of the position of an active region.

The NSA Kalman filter is a variant of the Kalman filter (Welch et al. 1995), which is more suitable for solving the problem of low-quality detection or noise during tracking. The NSA Kalman filter includes two main stages, prediction and update.

The prediction stage includes state prediction and covariance prediction. The state prediction predicts the position and movement information of the target $\hat{x}_k^-$ according to the state transition matrix F and the prior state estimate $\hat{x}_{k-1}^-$. The $\hat{x}_k^-$ is calculated as follows:

$$\hat{x}_k^- = F \cdot \hat{x}_{k-1}^-, \tag{1}$$

where the state transition matrix $F = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}$ is a 2×2 matrix, Δt is the time interval between two moments, and $\hat{x}_{k-1}^-$ represents the state of the target in the previous frame.

The covariance prediction adopts the prior covariance matrix and the process noise covariance matrix to predict the estimated covariance matrix P_k^- as follows:

$$P_k^- = F \cdot P_{k-1} \cdot F^T + Q, \tag{2}$$

where P_{k-1} is the covariance matrix in the $(k-1)$th frame, which is a symmetric semidefinite matrix whose diagonal elements represent the variance of the estimate. This matrix represents the difference, or error, between the estimated state of the real system and the true state. The diagonal elements provide a measure of uncertainty for each state estimate, where smaller diagonal elements indicate that the estimate is more accurate, and vice versa. The process noise covariance matrix Q describes the uncertainties introduced by external factors.

To improve the accuracy of the predicting positions of active regions, we revised the prior state estimate $\hat{x}_{k-1}^-$ according to the differential rotation theorem (Welch et al. 1995) in the prediction stage. The differential rotation theorem is described as

$$\omega = A + B \sin^2(\phi) + C \sin^4(\phi), \tag{3}$$

where ω is the rotational angular velocity, and ϕ is the solar latitude. The constants A, B, and C are approximate values obtained by the least square method as $A = 2.894 \pm 0.011\ \mu\mathrm{rad\,s}^{-1}$, $B = -0.428 \pm 0.070\ \mu\mathrm{rad\,s}^{-1}$, and $C = -0.370 \pm 0.077\ \mu\mathrm{rad\,s}^{-1}$, respectively. Given ω and the time difference between the two frames, the horizontal position of the active region in the next frame can be estimated approximately.

In the update stage the prediction results will be revised by the detection results in the current frame. The update stage includes state estimation revision and covariance matrix revision. The Kalman gain K_k and the measurement residuals y_k need to be calculated as follows:

$$K_k = P_k^- \cdot H^T \cdot (H \cdot P_k^- \cdot H^T + R_k)^{-1}, \tag{4}$$

$$y_k = z_k - H \cdot \hat{x}_k^-, \tag{5}$$

where P_k^- is the estimated covariance matrix, R_k is the noise covariance matrix, z_k is a vector including the position and velocity of the target estimated by differential rotation theorem at time k, and H is the measurement matrix, which represents the output of the detection module, containing the position, size, shape, and other information of targets.

The noise covariance $\tilde{R}_k$ is adaptively calculated for improving the accuracy of the updated state as follows:
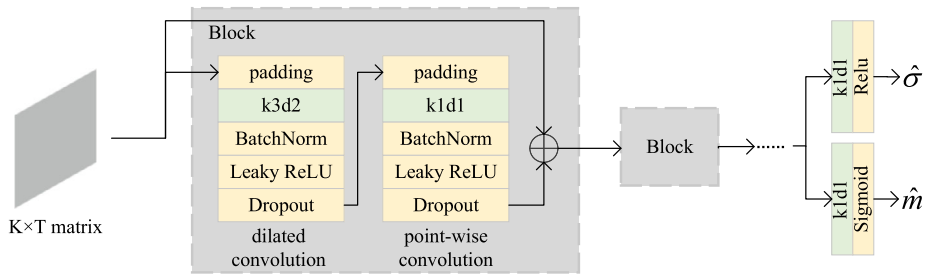
$$\tilde{R}_k = (1 - c_k)\, R_k, \tag{6}$$

Figure 2 The structure of the splitter module.

where R_k is the measurement noise covariance, and c_k is the detection confidence score at state k. The score reflects the level of noise, with higher score being less noisy, resulting in lower $\tilde{R}_k$. A lower $\tilde{R}_k$ means that the detection has a higher weight in status update stage, and vice versa.

The state estimation revision mainly measures the updated state estimate $\hat{x}_k$, which is calculated by mapping the measurement residual into the state estimate using Kalman gain:

$$\hat{x}_k = \hat{x}_k^- + K_k \cdot y_k. \tag{7}$$

The covariance matrix revision measures the covariance matrix P_k, which is also modified by Kalman gain:

$$P_k = (I - K_k \cdot H) \cdot P_k^-. \tag{8}$$

So far, the position information and motion information of the active regions have been initially predicted. The motion information will be further revised in the splitter module. The position information will be used in the AR association module.

3.2.2. Splitter Module

To solve the trajectory interruption problem caused by the deformation and temporary disappearance of active regions, we integrated the splitter module (Wang et al. 2022) for predicting ID-change position of each active region. The input of the module is a matrix of dimensions $K \times T$, where K is the number of targets, and T is the temporal length of the trajectories. Each element in the matrix represents a feature value of a target that incorporates motion information at that time.

The structure of the splitter module is shown in Figure 2. It stacks 16 temporal dilated convolution blocks, each block mainly including a dilated convolution and a pointwise convolution. The dilated convolution in sequence includes a padding layer, a dilation convolution, a Batchnorm layer, Leaky ReLU activation function (Glorot, Bordes, and Bengio 2011), and a dropout layer. The dilation convolution adopts a dilation rate for expanding the receptive field by inserting interval zero values into the convolution kernel. For example, the label $k3d2$ means the temporal dilation convolution with kernel size 3 and dilation rate 2. The pointwise convolution is similar, which uses a 1×1 convolution for reducing the channels to decrease the calculation and the parameters. To capture temporal patterns, a skip connection in each intermediate block is added, except the first and last blocks. The skip connection directly connects the features of an active region in the current block with that in

the previous block. After 16 blocks, the features of an active region have captured long-term temporal information. Following a 1×1 convolution, a one-dimensional vector $\hat{z}$ reflects the difference between active regions in the contiguous frames. They are performed by Sigmoid and ReLu activation functions for obtaining the ID-change position mask $\hat{m}_t$ and standard deviation σ, respectively. The ID-change position mask m_t is defined as follows:

$$m_t = \begin{cases} 1 & \text{if } sigm\left(\hat{z}\right) > T, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

The ID-change position value in the frame t is set to 1 if the $sigm\left(\hat{z}\right)$ value is larger than the threshold T, indicating an occurrence of ID-change. Here T is set to 0.7 after experiments. The ID-change position mask will guide the forward and backward association of the tracking target in the next association module. The standard deviation is a time and trajectory-related value used to adjust the smoothness of the loss function. The larger the standard deviation, the higher the probability of ID-change at that location.

The target-ID embeddings in consecutive frames usually change gradually rather than change abruptly, even in the ID-change positions. Therefore soft labels are adopted, which are continuous and smooth representation of labels that provide more uncertainty. The loss function is designed as the difference between the predicted ID-change position mask and the real ID-change position mask by an adaptive Gaussian smoothing strategy. The smoothness of the soft label is controlled by the standard deviation σ, which is adaptively adjusted according to the duration of the ID-change. A larger σ value will make the soft label smoother, and vice versa. The loss function is as follows:

$$L_s = \sum_t \left(\hat{m}_t - \min\left(\sum_\tau m_\tau^* \exp\left(-\frac{(\tau - t)^2}{\sigma^2} \right), 1 \right) \right), \tag{10}$$

where t and τ represent the frame index of the predicted soft label and the ground truth label, respectively, and $\hat{m}_t$ and m_τ^* represent the ID-change position values of predicted soft labels and ground truth labels at frame index t and τ, respectively.

3.2.3. AR Association Module

The AR association module is designed to associate active regions between consecutive frames by a flow graph. As shown in Figure 1, the detection_all coming from detection module, the position information coming from NSA Kalman filter, and ID-change position mask coming from the splitter module are concatenated as input.

The detailed structure is shown in Figure 3. The input is feature extracted by the Resnet-18 network firstly. The extracted features are being nodes to build the flow graph, in which each edge connecting two nodes represents the matching probability of these nodes in the consecutive frames. The target initial state set S includes all targets consisting of the target-ID and motion information in all frames. The Euclidean distance and IOU value between any two targets in the consecutive frames are calculated according to the information saved in S. The initial flow graph is constructed as each node in the first frame is added an edge to the node in the second frame with the shortest Euclidean distance and IOU value greater than 0.5. This process repeats from the second frame until the last frame. Meanwhile, G is completed as a set of all edges corresponding the flow graph. Then the flow graph is fed into a multilayer perceptron (MLP) to calculate the connection probability of each edge according to the features of corresponding nodes. Each node only reserves two edges with
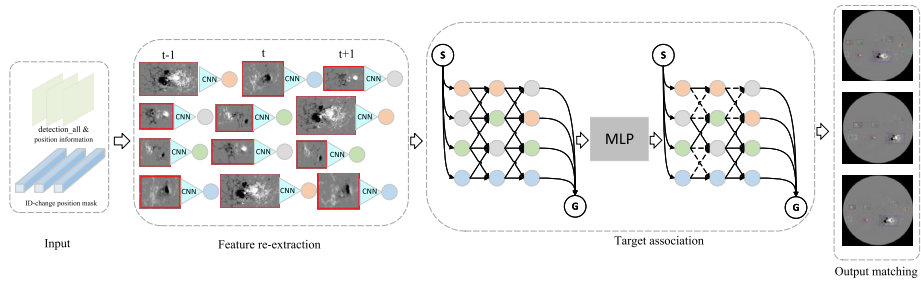
Figure 3 The structure of AR association module.

the greatest probability of the previous connection and next connection. The set G updates simultaneously.

Using this module, the multitarget tracking problem is transformed into the minimum-cost network flow problem. The connection probability is inferenced by the negative log likelihood function, which corresponds a constrained integer linear programming problem. The optimal association results could be obtained, resulting in the target tracking.

In this module, MLP is used to calculate the cost of edges in the flow graph by learning the parameters and ID-change position mask using a cost function. The minimum cost flow problem is transformed into a constrained integer linear programming as follows:

$$L_a = -\sum_i y_i \cdot \log(p_i), \tag{11}$$

where y_i represents the ground truth in frame i, and p_i is the predicted value of the module output.

To optimize the parameters iteratively through training, a multitask loss is adopted. The loss function of the whole model is defined as

$$L_{\text{total}} = \frac{1}{2} \left(\frac{1}{e^{w_1}} \left(L_{\text{heat}} + L_{\text{box}} \right) + \frac{1}{e^{w_2}} \left(L_{\text{s}} + L_{\text{a}} \right) + w_1 + w_2 \right), \tag{12}$$

where w_1 and w_2 are the parameters that balance the detection module and tracking module. Here w_1 and w_2 were set as 0.5. The loss function of detection module is defined as the sum of L_{heat} and L_{box}; L_{heat} represents the heatmap loss defined as pixelwise logistic regression with focal loss (Lin et al. 2017b), and L_{box} represents the box offset and size loss, which is defined as the sum of two l_1 losses; L_{s} and L_{a} are the loss function of splitter module and AR association module defined by formulae 10 and 11, respectively.

Summarily, most active regions are detected from solar images in sequences through base detection module first, and then a few undistinguished active regions that are missed in the base detection module are captured in the re-detection module. The above detected active regions are integrated and then transferred to tracking step. These detected active regions are predicted by their position information and motion information by NSA Kalman filter firstly, and then ID-change position of each active region is predicted by the splitter module for solving the trajectory interruption problem. Finally, these amended active regions are associated between successive frames by a flow graph in the AR association module. So far, the task of detecting and tracking active regions in solar magnetogram sequences is finished.

4. Experiments

4.1. Training and Testing

This deep-learning method, ReDetGraphTracker, as an end-to-end system, which integrates all the modules detailed in Section 3 as a pipe-line. The method was deployed on a personal computer, equipped with one Nvidia GeForce 2080 GPU. The main environment includes Ubuntu 16.04, CUDA 10.2, PyTorch 1.7.0, Python 3.7, and opencv-python 4.0.0.21.

Training it only needs to set the path of training set. To achieve optimal tracking performance, the optimizer was selected as Adam. The batch size was set to 8. The model was trained over 200 epochs with a learning rate of $1e^{-4}$, which takes about 46 hours until the loss value converges stably. By the way, the redetection network incorporates an auxiliary training module (ATB) to expedite model convergence. During training, the features produced by the backbone are concurrently fed into both the detection module and ATB. Since ATB has a larger parameter space, it is easier to learn the initial state distinction of positive and negative sample partition and label matching. Using ATB to guide the tag matching makes the detection branch converge faster.

After training the method, the method was tested on all testing sets by setting their corresponding testing paths. It took about one second to process an image on average, meeting near-real-time requirements. The results will be detailed in the next section. The code is publicly available at github.com/Caesarisis/ReDetGraphTracker.

So far, once this trained method is deployed, any image series consisting of sequential magnetograms are set as inference set, and the end-to-end system can detect and track the active regions in near-real-time mode. It will output the sequential magnetograms labeled with tracking bounding boxes and a text file including the track-ID and the coordinates of bounding boxes.

4.2. Metrics

The performance evaluation uses typical multiobject tracking metrics (Bernardin and Stiefelhagen 2008), including multiobject tracking accuracy (*MOTA*), multiobject tracking precision (*MOTP*), and identity switching (*IDs*). *MOTA* represents the overall performance of the model in multitarget tracking, defined as

$$MOTA = 1 - \frac{FN + FP + IDs}{GT},$$ (13)

where *FN* is the number of positive samples predicted by the model to be in the negative class, and *FP* is false positive and represents the number of classes that are incorrectly labeled as positive among the negative samples. *IDs* is the total number of ID switches, or the number of target mismatches. *GT* is the number of Ground Truth.

MOTP represents the average overlap between all tracked targets defined as

$$MOTP = \frac{\sum_{t,i} d_{t,i}}{\sum_t c_t},$$ (14)

where c_t represents the total matching degree between the ground truth and the detection output at frame index t, and $d_{t,i}$ represents the overlap of the bounding box i between the target and the ground truth at frame index t.

Table 3 Tracking metrics.

Dataset	IDF1↑	MOTA↑	MOTP↓	IDs↓	FPS↑
D_1	75.4%	75.3%	0.128	16	13.6
D_2	73.2%	74.1%	0.130	14	14.0
D_3	76.5%	77.4%	0.124	16	13.8
D_4	73.1%	74.1%	0.132	11	13.2
D_5	73.4%	73.8%	0.129	13	13.3
D_6	72.1%	73.2%	0.137	12	13.4
D_{6-1}	82.5%	87.3%	0.120	11	13.2
D_{6-2}	86.5%	88.3%	0.113	8	13.1
D_{6-3}	88.3%	91.2%	0.102	3	13.3
D_{6-4}	60.4%	57.3%	0.234	58	13.5

IDF1 is a performance evaluation of object associations. It is based on the number of false positive identities (*IDFP*), false negative identities (*IDFN*), and true positive identities (*IDTP*). *IDF1* defined as

$$IDF1 = \frac{2IDTP}{2IDTP + IDFP + IDFN}. \tag{15}$$

This metric is an identity-based evaluation, which allows for a more comprehensive assessment of tracking performance because it takes into account any potential changes in object appearance or motion over time. For example, an object that is tracked correctly in one frame but not in the next frame will be identified as an error in the identity-based evaluation, which will not be captured in a frame-by-frame comparison.

5. Results

5.1. Metrics and Instances

Table 3 lists the testing metrics of the ten testing sets. All the metrics of the testing sets with interval less than 24 h ($D_{6-1} \sim D_{6-3}$) are much better than those of the testing sets with 24-h interval ($D_1 \sim D_6$) as expected, whereas the testing set with 48-h interval (D_{6-4}) does not work so well. We think the method is feasible for the image sequence within 24 hours. The averaged metrics *IDF1*, *MOTA*, *MOTP*, *IDs*, and *FPS* for six testing sets with 24-h interval are 74.3%, 75.0%, 0.130, 14.5, and 13.7, respectively. The metrics of HMI and MDI data are similar. This means that the method has a good generalization and will well process the solar image sequences coming from the other telescopes. The metrics get worse with the number of active regions increasing, e.g., D_6. The complexity of targets will increase the difficulty of detection and tracking. With the intervals decreasing, the model has an excellent perform, e.g., for D_{6-3} with 12-min interval, *MOTA* is up to 91.2%, and *IDs* is only 3. The feature of active region keeps without much change because it evolves slowly in a short time. So we believe that the key to get good results is the interval of image sequences, and the second is the complexity of solar activities.

Figure 4 shows a tracking segment of four consecutive frames in D_5. The label at the top of each active region is labeled automatically by this method, whereas the label at the
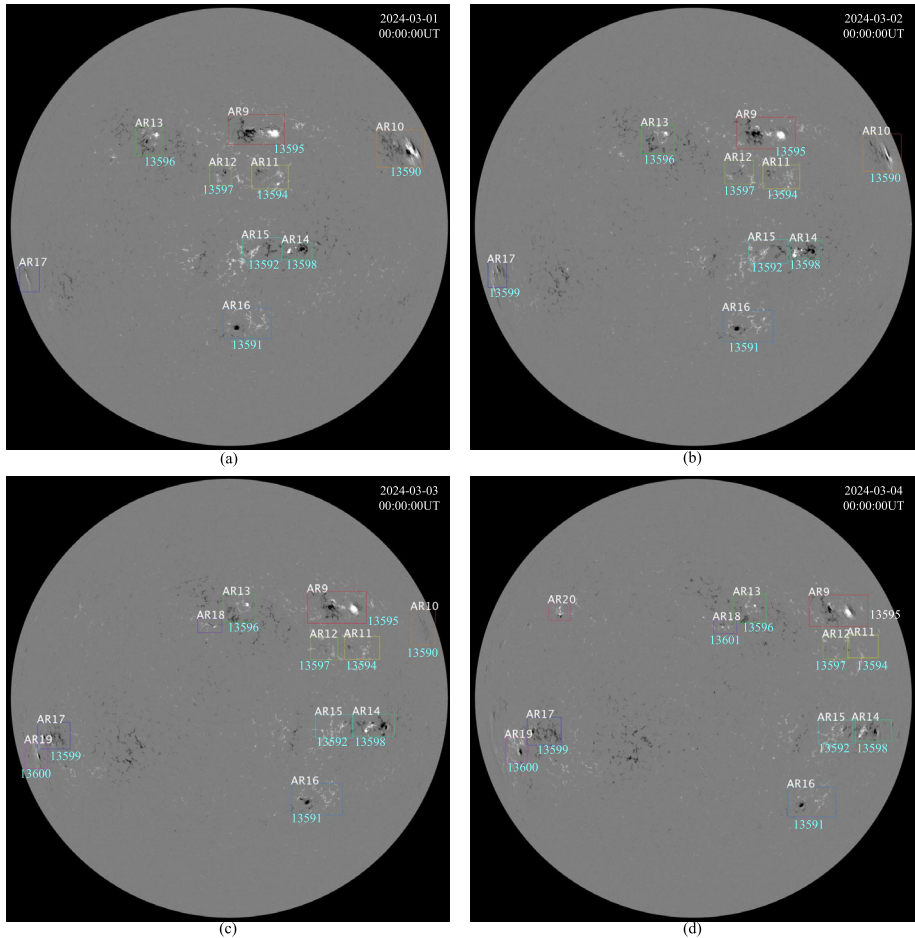
Figure 4 The tracking results of ReDetGraphTracker for four consecutive frames from 1 March 2024 to 4 March 2024. The label at the top of each active region is labeled automatically by this method, whereas the label at the bottom is artificial marked according to NOAA for comparison.

bottom is artificial marked according to NOAA for comparison. Most active regions by two methods are the same, especially those active regions with obvious characteristics. The main differences come from those active regions appearing at the edge of Sun, or the weak active regions. For instance, AR17 is detected and tracked on 1 March 2024 at once, while it appears at the solar limb. The corresponding NOAA13599 is labeled on 2 March 2024, which is later one day. AR20 is detected and tracked on 4 March 2024, and the corresponding NOAA13603 will be labeled until 6 March 2024. The ReDetGraphTracker can detect and track the active region in near-real-time once the active region shows visible magnetic fields.

Figures 5 (a) and (b) show the complete tracking results of the NOAA11087 in the D_2 (MDI) and D_3 (HMI), respectively. The bounding boxes obtained by ReDetGraphTracker are drawn with green solid line, labeled as AR5 of MDI data and AR6 of HMI data, respectively. The results coming from SMARPs and SHARPs are drawn with yellow dashed line

(a) SHARPs&ReDetGraphTracker

(b) SMARPs&ReDetGraphTracker
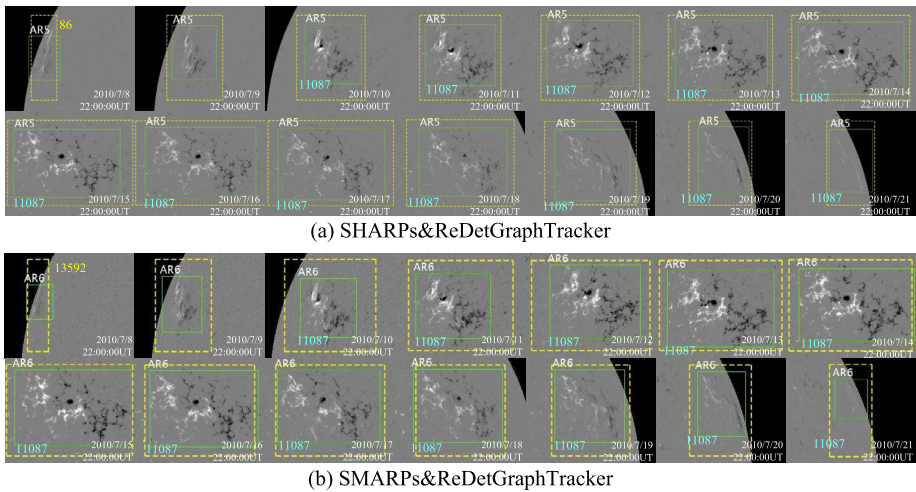
Figure 5 Panels (a) and (b) show the complete tracking results of NOAA11087 in the D_1 (MDI) and D_3 (HMI), respectively. A total of 16 frames are displayed. The bounding boxes obtained by the ReDetGraph-Tracker are drawn with green solid line, labeled as AR5 of MDI data and AR6 of HMI data, respectively. The results coming from SMARPs and SHARPs are drawn with yellow dashed line with number 86 in Panel (a) and with number 13592 in Panel (b), respectively. Also, the label at the bottom is artificial marked according to NOAA.

with number 86 in Panel (a) and with number 13592 in Panel (b), respectively. Also, the label at the bottom is artificial marked according to NOAA. By the way, there is no NRT SHARPs in this period of time. The bounding boxes of SMARPs and SHARPs enclose the same heliographic area. We can see that the ReDetGraphTracker obtains relatively stable bounding boxes in a near-real-time mode, whatever the MDI data or HMI data. The active region is labeled from 8 July by SMARPs, SHARPs, and ReDetGraphTracker, whereas it is labeled from 10 July by NOAA. The ReDetGraphTracker method captures it in a near-real-time, tracks it with suitable bounding box continuously, and finally terminates it in time.

Figures 6 (a) and (b) show the total line-of-sight unsigned magnetic flux and the pixel area mapping to a heliographic Cylindrical Equal-Area coordinate system over time for NOAA11087 as calculated from LoS magnetic field maps. Since only the bounding box of active regions achieved (without mask) by the ReDetGraphTracker, the magnetically active pixels in each bounding box in HMI data are approximately determined with absolute LoS field greater than 100 G (Hoeksema et al. 2014). After reviewing the relevant literature, the magnetically active pixels in each bounding box in MDI data are approximately determined with absolute LoS field greater than 56 G (Bobra et al. 2021). The corresponding data from SMARPs and SHARPs are plotted for comparison. The ReDetGraphTracker method achieved consistent total unsigned magnetic flux results with both SHARP data on HMI and SMARP data on MDI.

In short, the ReDetGraphTracker method has a good performance in tracking active regions in solar full-disk LoS magnetograms, especially obtaining a relatively accurate life-time from its appearance to disappearance. The active regions whatever evolve drastically or with weak magnetic field strengths will be detected and tracked effectively in near-real-time mode.
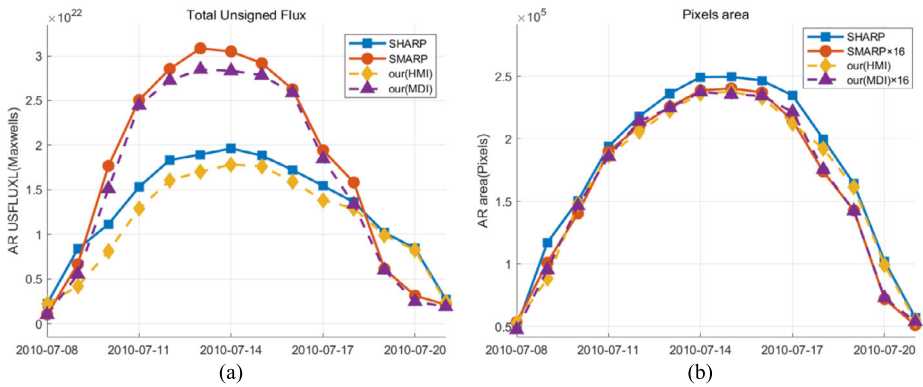
Figure 6 Panels (a) and (b) show the total line-of-sight unsigned magnetic flux and the pixel area mapping to a heliographic Cylindrical Equal-Area coordinate system over time for NOAA11087 as calculated from LoS magnetic field maps. The magnetically active pixels in each bounding box are approximately determined with absolute LoS field greater than 100 G (Hoeksema et al. 2014). The corresponding data from SMARPs and SHARPs are plotted for comparison (the data also can be seen in (Bobra et al. 2021, Figure 2)). The data of ReDetGraphTracker are very consistent with those of SMARPs and SHARPs.

Table 4 The metrics of ablation experiments on D_4.

Models	IDF1↑	MOTA↑	MOTP↓	IDs↓	FPS↑
BM	52.7%	53.4%	0.195	48	25.2
BM+M1	55.4%	58.9%	0.167	45	21.8
BM+M2	63.3%	60.4%	0.159	36	23.2
BM+M1+M2	66.1%	65.3%	0.149	34	20.6
BM+M2+M3	69.4%	67.4%	0.144	16	16.2
BM+M1+M2+M3	72.1%	73.2%	0.137	12	13.4

5.2. Ablation Experiments

The basic tracking model (BM) is taken as performing detection and tracking task basically, which is an integration of base detection module and AR association module in this work. Based on it, the other modules were combined gradually to test their necessity and feasibility, such as redetection module representing as M1, NSA Kalman filter as M2, and the splitter module as M3. Since M3 requires data from M2, M3 needs to be tested integrated with M2 in this ablation experiments.

This ablation experiment was verified on the D_4 data set. The metrics are shown in Table 4. The base model (BM) starts with 58.2% IDF1, 53.4% MOTA, 0.195 MOTP, 48 IDs, and 25.2 FPS, respectively.

After adding M1 to the base model, most metrics improve, except FPS. IDF1 and MOTA increase by 2.7% and 5.5%, respectively. The MOTP decreased by 0.028. Both the differences of IDs and IDF1 indicate that the M1 module has improved on the maintenance continuous trajectory. Additionally, this module increases precision while simultaneously reducing computational complexity, which make FPS only reduce 3.4.

Only adding M2 to the base model, most metrics further improve, especially IDF1. Examples of BM+M1 and BM+M2 are shown in Figures 7 (a) and (b), respectively. We can
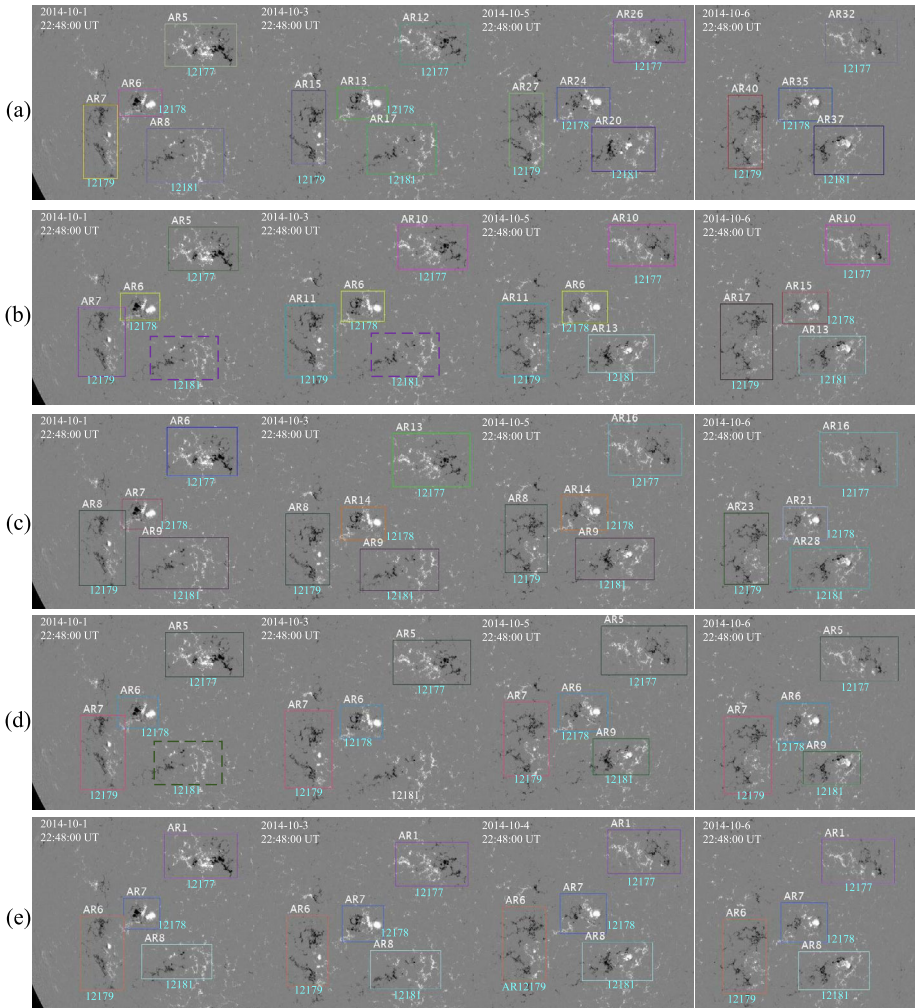
Figure 7 The comparison of ablation experiments. The basic tracking model (BM) is an integration of base detection module and AR association module. The redetection module is represented as M1, NSA Kalman filter as M2, and the splitter module as M3. Panel (**a**) is the result of BM+M1, (**b**) is the result of BM+M2, (**c**) is the result of BM+M1+M2, (**d**) is the result of BM+M2+M3, and (**e**) is the result of BM+M1+M2+M3 or the final model, ReDetGraphTracker.

see that more active regions are detected in Figure 7 (a), owing to the redetection module, e.g., AR8 (NOAA12181) in the first column and AR17 (NOAA12181) in the second column. The more reliable track-IDs are kept, e.g., AR6 (NOAA12178) in the second and third columns.

With M1, M2, and M3 combining to the BM using different combinations, both *IDF1* and *MOTA* increase gradually, whereas the *MOTP* decreases gradually. The continuous decrease in *IDs* indicates that these modules help track-ID changes reduced. The metrics of different combinations show that all these modules are independent and useful. The less *FPS* drop

Table 5 The metrics of ReDetGraphTracker and other two methods on D_4.

Models	IDF1↑	MOTA↑	MOTP↓	IDs↓	FPS↑
DeepSort	56.3%	61.2%	0.225	65	38.3
CenterTrack	67.4%	66.3%	0.209	49	22.8
SolarTracking	68.8%	68.1%	0.189	31	20.5
ReDetGraphTracker (ours)	72.1%	73.2%	0.137	12	13.4

of BM+M2 indicates that the M2 is a lightweight module. Even at the final integration, the *FPS* drops only a half.

The result of BM+M1+M2 is shown in Figure 7 (c). In the subregions from 1 October 2014 to 6 October 2014, BM+M1+M2 can well detect the active regions. Compared with NOAA, there is no misdetection. Only adding M1 and M2 cannot solve the problem of active region ID change well. For example, NOAA12178 had four ID changes during six-day period, and NOAA12179 and 12181 also had an ID change on 6 October 2014.

The detection and tracking result of BM+M2+M3 is shown in Figure 7 (d). The problem of ID change is well solved. However, NOAA12181 is not detected until 5 October 2014 by BM+M2+M3.

The result of BM+M1+M2+M3 is shown in Figure 7 (e). We can see that there is no missing detection and no problem of ID change in the final model, ReDetGraphTracker.

In short, the redetection module M1 plays a significant role in improving the detection results. The NSA Kalman filter M2 plays a key role in the continuity of the trajectory and the stability of the tracking bounding box. M1 and M2 are relatively independent. The splitter module M3 has good effect on solving the problem of ID change, which is dependent on M2.

5.3. Comparison

We also compared ReDetGraphTracker with other typical methods, including two general tracking methods DeepSORT (Wojke, Bewley, and Paulus 2017) and CenterTrack (Zhou, Koltun, and Krähenbühl 2020), one tracking method for solar events (SolarTracking; Kempton, Pillai, and Angryk 2014), NRT SHARPs, and SHARPs. DeepSORT uses Kalman filtering to predict the location of the bounding box and the Hungarian algorithm to match the current detection with the largest IOU value. CenterTrack calculates the distance matrix between the center point of targets in the consecutive frames and then selects the target with the smallest Euclidean distance for matching. SolarTracking predicts the location based on the differential rotation theorem and then performs target association based on Euclidean distance. For comparison, the detection modules of all these three methods were assembled as the base detection module in ReDetGraphTracker method. The hyperparameter settings for training them were consistent with ReDetGraphTracker method. All the methods were tested by D_4. Table 5 lists the evaluation metrics of DeepSORT, CenterTrack, SolarTracking, and ReDetGraphTracker. The metrics show that ReDetGraphTracker obtains good tracking effects, especially in *IDs*. Due to combined multiple modules, the tracking speed is slow.

Figure 8 shows an example corresponding to NOAA13627 from 1 April 2024 to 14 April 2024 for comparing these tracking methods. Panel (a) is the tracking result of the DeepSORT. Panel (b) is of the CenterTrack. Panel (c) is of the SolarTracking. Panel (d) is of our ReDetGraphTracker, and the results of the NRT SHARPs and SHARPs.

NOAA13627 is labeled by NOAA from 3 to 14 April. Using DeepSORT, it is tracked from 3 to 11 April. The unsatisfactory thing is that it occurs five ID changes on 4, 6, 7,
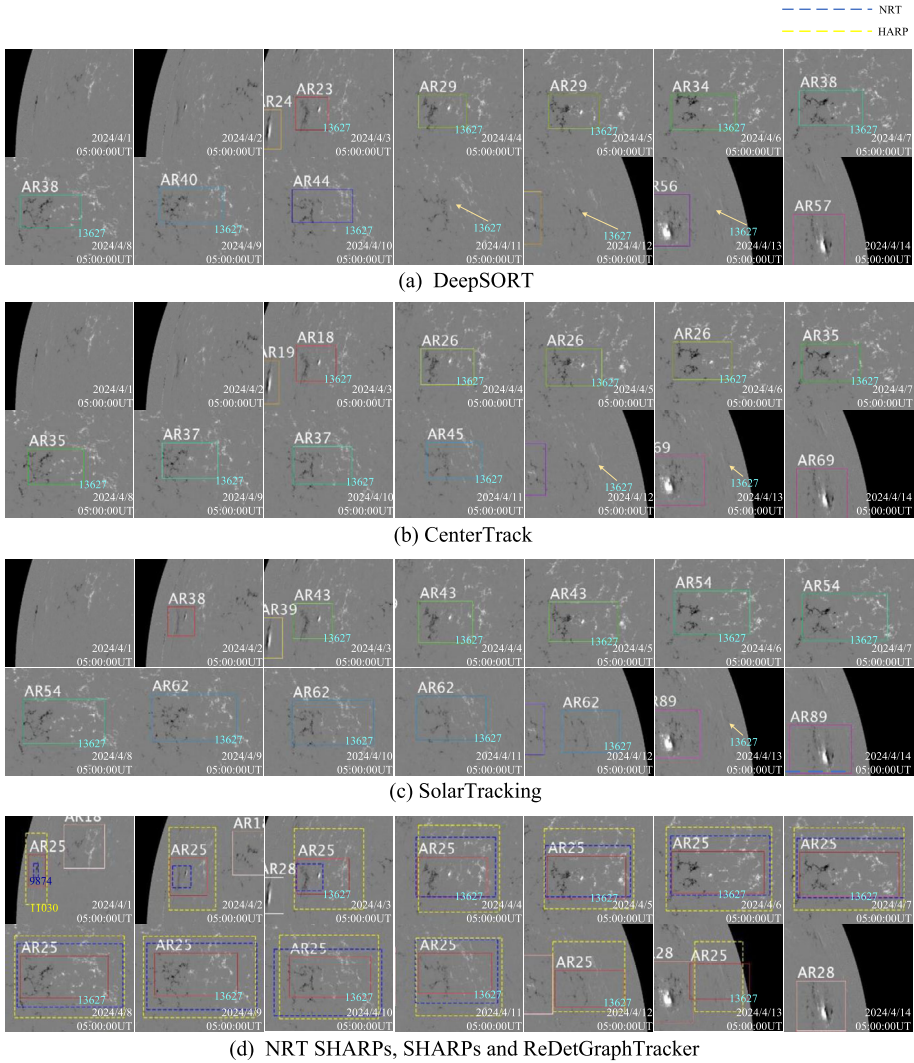
Figure 8 An example corresponding to NOAA13627 from 1 April 2024 to 14 April 2024 for comparing these tracking methods. There are a total of 14 frames. The label at the top of each active region is labeled automatically by the corresponding method, whereas the label at the bottom is artificially marked according to NOAA for comparison. Panel (**a**) is the tracking result of the DeepSORT (Wojke, Bewley, and Paulus 2017). Panel (**b**) is of the CenterTrack (Zhou, Koltun, and Krähenbühl 2020). Panel (**c**) is of the SolarTracking (Kempton, Pillai, and Angryk 2014). Panel (**d**) is of the ReDetGraphTracker (AR25, with red solid box), and the results of the NRT SHARPs (9874, with blue dashed box) and SHARPs (11030, with yellow dashed box). The red arrows point to the regions labeled with NOAA but not tracked by the corresponding method.

9, and 10 April with a large error bounding box (see Panel (a)). Using CenterTrack, it is terminated tracking on 12 April. Additionally, it occurs four ID changes on 4, 7, 9, and 11 April (see Panel (b)). Using SolarTrack, it is tracked from 2 to 13 April. Three ID changes occur on 3, 6, and 9 April (see Panel (c)). Compared with DeepSORT and CenterTrack, SolarTrack obtains relatively good bounding boxes.
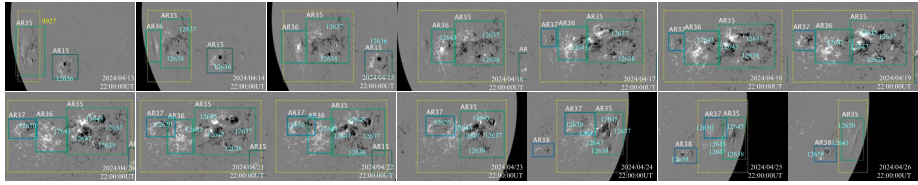
Figure 9 This shows an example including very active regions from 14 April 2024 to 25 April 2024. There are a total of 14 frames. The results of the ReDetGraphTracker are drawn with solid box, and of the SHARPS with yellow dashed box. The NOAA numbers are labeled at the region in light-blue.

Panel (d) shows the results of the ReDetGraphTracker (AR25, with red solid box), the results of the NRT SHARPs (9874, with blue dashed box), and SHARPs (11030, with yellow dashed box). All of them are labeled from 1 April, whereas it is labeled from 3 April by NOAA. Since the NRT SHARPs is processed within several hours after observation, its heliographic size may change during its life, e.g., on 4 April. Compared with NRT SHARPs, the bounding box of a definitive HARP encloses the same heliographic area during its entire lifetime when its whole history has been known. The ReDetGraphTracker obtains relatively stable bounding boxes in a near-real-time mode. At last, tracking terminated on 12 April by NRT SHARPs, whereas it terminated on 14 April by NOAA, SHARPs, and ReDetGraphTracker.

Figure 9 shows an example including very active regions from 14 April 2024 to 25 April 2024. The results of the ReDetGraphTracker are drawn with solid box, and of the SHARPs with yellow dashed box. The NOAA numbers are labeled at the region in light-blue. We can see that it is a very active region, which includes six NOAA active regions. Since the definitive SHARPs processing module groups and tailors the identified regions according to their complete life history, all of these active regions are grouped as a patch (9927). In a near-real-time mode, the ReDetGraphTracker processes every active region in each image frame one by one. There are three active regions detected and tracked (AR35, AR36, and AR37), which keep stable bounding boxes and track-ID. This means that if the active regions are far from others, then the ReDetGraphTracker can obtain the results similar to those of NOAA; e.g., see Figure 4. Otherwise, a region may contain multiple NOAA ARs.

In short, ReDetGraphTracker integrates base detection network, redetection network, NSA Kalman filter, the splitter module, and a flow graph in the AR association module. It performs well in detecting and tracking active regions, at a cost of speed. Importantly, it will obtain relative stable detection and tracking of active regions in a near-real-time mode.

6. Conclusions

The solar active region is the region with strong magnetic fields, which is the main energy source of solar activities. Violent solar activities lead to severe space weather, such as flares and coronal mass ejections (CME), which have adverse effects on the Earth, e.g., electromagnetic telecommunication, electric power system, radio transmissions, etc.

A total of 4577 HMI and MDI magnetograms from 12 July 2000 to 15 May 2024 are selected for building the data set. Both the training and validating sets come from HMI, and the testing set comes from HMI and MDI. The time intervals were set as 12 minutes, 1 hour, 6 hours, 24 hours, and 48 hours. The testing set consists of ten sets of image sequences, which represent different observation instruments, different numbers of activity regions, and different time intervals.

This study presents a new deep learning method, ReDetGraphTracker, for detecting and tracking the active regions in full-disk longitudinal magnetograms. Firstly, most obvious active regions are detected through base detection network; secondly, those undistinguished active regions that are missed through the base detection network are captured in the redetection network; thirdly, these detected active regions are predicted their position and motion information by NSA Kalman filter combining with differential rotation theory; fourthly, ID-change position of each active region is predicted by the splitter module for reducing the trajectory interruption; and finally, these amended active regions are associated between successive frames by a flow graph in the AR association module.

The model was trained 46 hours for 200 epochs until the value of loss converged stably on a personal computer with one Nvidia GeForce 2080 GPU. After training, ten testing sets were fed into the model to test. The evaluation metrics *IDF1*, *MOTA*, *MOTP*, *IDs*, and *FPS* for the testing sets with 24-h interval on average are 74.0%, 74.7%, 0.130, 13.6, and 13.6, respectively. With the decreasing intervals, the metrics improve gradually. For the testing set with an interval of 12 minutes, it achieves the best evaluation, where *IDF1*, *MOTA*, *MOTP*, *IDs* and *FPS* are 88.3%, 91.2%, 0.102, 3, and 13.3, respectively. The method is not feasible for the image sequence with time interval beyond 24 hours. Additionally, it has good generalization.

We analyzed the instances obtained by ReDetGraphTracker in detail and compared with the other typical tracking methods, DeepSORT, CenterTrack, SolarTracking, SMARPs, SHARPs, and NRT SHARPs. Additionally, each module of ReDetGraphTracker was validated by ablation experiments. All the above results show that ReDetGraphTracker has a good performance in detecting and tracking active regions in full-disk LoS magnetograms. The main reason comes from these cooperative modules, especially the redetection module, NSA Kalman filter, and the splitter module. The redetection module plays a significant role in improving the detections, especially those active regions with weak features. The NSA Kalman filter, combined with differential rotation theory, plays a key role in the continuity of the trajectory and the stability of the tracking bounding box. The splitter module has good effect on solving the problem of ID change. Benefiting from these modules, the ReDetGraphTracker can capture an active region as early as possible and terminate tracking in near-real time. Additionally, the active regions, regardless of whether they evolve drastically or have weak magnetic field strengths, will be detected and tracked well. The last advantage of this method is near-real time. It detects and tracks the active regions for an image sequence in time, which do not need to know the whole history of active regions. In the future the method can try to accept an image one by one to achieve stronger near-real-time tracking.

Data Availability No datasets were generated or analysed during the current study.

Materials Availability The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request. The movies of all the testing sets are public at github.com/Caesarisis/ReDetGraphTracker.

Declarations

Competing interests The authors declare no competing interests.

References

Bai, H., Yang, P., Zhao, L., Gong, X., Zhong, L., Yang, Y., Rao, C.: 2023, Hybrid detection algorithm and study on the quantity and brightness evolution characteristics of photospheric bright point groups. *Astrophys. J.* **956**, 62.

Bernardin, K., Stiefelhagen, R.: 2008, Evaluating multiple object tracking performance: the CLEAR MOT metrics. *EURASIP J. Image Video Process.* **2008**, 1.

Bobra, M.G., Wright, P.J., Sun, X., Turmon, M.J.: 2021, Smarps and sharps: two solar cycles of active region data. *Astron. Astrophys. Suppl. Ser.* **256**, 26.

Caballero, C., Aranda, M.: 2014, Automatic tracking of active regions and detection of solar flares in solar EUV images. *Solar Phys.* **289**, 1643.

Colak, T., Qahwaji, R.: 2009, Automated solar activity prediction: a hybrid computer platform using machine learning and solar imaging for automated prediction of solar flares. *Space Weather* **7**.

darkpgmr: 2020, DarkLabel, https://github.com/darkpgmr/DarkLabel.git.

Du, Y., Wan, J., Zhao, Y., Zhang, B., Tong, Z., Dong, J.: 2021, GIAOTracker: a comprehensive framework for MCMOT with global information and optimizing strategies in VisDrone 2021. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2809.

Gallagher, P.T., Moon, Y.-J., Wang, H.: 2002, Active-region monitoring and flare forecasting – I. Data processing and first results. *Solar Phys.* **209**, 171.

Giersch, O., Kennewell, J., Lynch, M.: 2018, Reanalysis of solar observing optical network sunspot areas. *Solar Phys.* **293**, 138.

Glorot, X., Bordes, A., Bengio, Y.: 2011, Deep sparse rectifier neural networks. In: *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, *JMLR Workshop and Conference Proceedings*, 315.

Guo, X., Yang, Y., Feng, S., Bai, X., Liang, B., Dai, W.: 2022, Solar-filament detection and classification based on deep learning. *Solar Phys.* **297**, 104.

He, K., Zhang, X., Ren, S., Sun, J.: 2016, Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770.

Higgins, P.A., Gallagher, P.T., McAteer, R.J., Bloomfield, D.S.: 2011, Solar magnetic feature detection and tracking for space weather monitoring. *Adv. Space Res.* **47**, 2105.

Hoeksema, J.T., Liu, Y., Hayashi, K., Sun, X., Schou, J., Couvidat, S., Norton, A., Bobra, M., Centeno, R., Leka, K., et al.: 2014, The Helioseismic and Magnetic Imager (HMI) vector magnetic field pipeline: overview and performance. *Solar Phys.* **289**, 3483.

Jiang, H., Jing, J., Wang, J., Liu, C., Li, Q., Xu, Y., Wang, J.T., Wang, H.: 2021, Tracing Hα fibrils through Bayesian deep learning. *Astron. Astrophys. Suppl. Ser.* **256**, 20.

Kempton, D., Pillai, K.G., Angryk, R.: 2014, Iterative refinement of multiple targets tracking of solar events. In: *2014 IEEE International Conference on Big Data (Big Data)*, IEEE, Washington, 36.

LaBonte, B., Georgoulis, M., Rust, D.: 2007, Survey of magnetic helicity injection in regions producing X-class flares. *Astrophys. J.* **671**, 955.

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: 2017a, Feature pyramid networks for object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2117.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P.: 2017b, Focal loss for dense object detection. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2980.

Liu, S., Qi, L., Qin, H., Shi, J., Jia, J.: 2018, Path aggregation network for instance segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 8759.

McAteer, R.J., Gallagher, P.T., Ireland, J., Young, C.A.: 2005, Automated boundary-extraction and region-growing techniques applied to solar magnetograms. *Solar Phys.* **228**, 55.

Meadows, P.: 2020, Remeasurement of Solar Observing Optical Network sunspot areas. *Mon. Not. Roy. Astron. Soc.* **497**, 1110.

Milan, A., Leal-Taixé, L., Reid, I., Roth, S., Schindler, K.: 2016, MOT16: a benchmark for multi-object tracking. arXiv preprint arXiv:

Mourato, A., Faria, J., Ventura, R.: 2024, Automatic sunspot detection through semantic and instance segmentation approaches. *Eng. Appl. Artif. Intell.* **129**, 107636.

Oludehinwa, I.A., Velichko, A., Belyaev, M., Olusola, O.I.: 2023, Imagery tracking of sun activity using 2D circular kernel time series transformation, entropy measures and machine learning approaches. arXiv preprint arXiv.

Qahwaji, R., Colak, T.: 2005, Automatic detection and verification of solar features. *Int. J. Imaging Syst. Technol.* **15**, 199.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: 2016, You only look once: unified, real-time object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779.

Scherrer, P.H., Schou, J., Bush, R., Kosovichev, A., Bogart, R., Hoeksema, J., Liu, Y., Duvall, T., Zhao, J., Title, A., et al.: 2012, The helioseismic and magnetic imager (HMI) investigation for the solar dynamics observatory (SDO). *Solar Phys.* **275**, 207.

Sreedevi, A., Jha, B.K., Karak, B.B., Banerjee, D.: 2023, AutoTAB: automatic tracking algorithm for bipolar magnetic regions. *Astron. Astrophys. Suppl. Ser.* **268**, 58.

Thompson, M.J., Christensen-Dalsgaard, J., Miesch, M.S., Toomre, J.: 2003, The internal rotation of the sun. *Annu. Rev. Astron. Astrophys.* **41**, 599.

Turmon, M., Pap, J., Mukhtar, S.: 2002, Statistical pattern recognition for labeling solar active regions: application to SOHO/MDI imagery. *Astrophys. J.* **568**, 396.

Turmon, M., Jones, H.P., Malanushenko, O.V., Pap, J.M.: 2010, Statistical feature recognition for multidimensional solar imagery. *Solar Phys.* **262**, 277.

Turmon, M., Hoeksema, T., Sun, X., Bobra, M.: 2011, HARPs: tracked active region patch data product from SDO/HMI. In: *2012 AGU Fall Meeting, and "TARPs: Tracked Active Region Patches from SoHO/MDI," 2013 AGU Fall Meeting*.

Wang, G., Wang, Y., Gu, R., Hu, W., Hwang, J.-N.: 2022, Split and connect: a universal tracklet booster for multi-object tracking. *IEEE Trans. Multimed.* **25**, 1256.

Welch, G., Bishop, G., et al.: 1995. An introduction to the Kalman filter.

Wojke, N., Bewley, A., Paulus, D.: 2017, Simple online and realtime tracking with a deep association metric. In: *2017 IEEE International Conference on Image Processing (ICIP)*, IEEE, Beijing, 3645.

Woo, S., Park, J., Lee, J.-Y., Kweon, I.S.: 2018, Cbam: convolutional block attention module. In: *Proceedings of the European Conference on Computer Vision (ECCV)*, 3.

Xu, L., Yang, Y., Yan, Y., Zhang, Y., Bai, X., Liang, B., Dai, W., Feng, S., Cao, W.: 2021, Research on multiwavelength isolated bright points based on deep learning. *Astrophys. J.* **911**, 32.

Yang, Y., Yang, H., Bai, X., Zhou, H., Feng, S., Liang, B.: 2018, Automatic detection of sunspots on full-disk solar images using the simulated annealing genetic method. *Publ. Astron. Soc. Pac.* **130**, 104503.

Yang, Y., Li, X., Bai, X., Zhou, H., Liang, B., Zhang, X., Feng, S.: 2019, Morphological classification of G-band bright points based on deep learning. *Astrophys. J.* **887**, 129.

Yu, F., Wang, D., Shelhamer, E., Darrell, T.: 2018, Deep layer aggregation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2403.

Zhang, J., Wang, Y., Liu, Y.: 2010, Statistical properties of solar active regions obtained from an automatic detection system and the computational biases. *Astrophys. J.* **723**, 1006.

Zharkova, V.V., Aboudarham, J., Zharkov, S., Ipson, S.S., Benkhalil, A.K., Fuller, N.: 2005, Solar feature catalogues in EGSO. *Solar Phys.* **228**, 361.

Zheng, Z., Hao, Q., Qiu, Y., Hong, J., Li, C., Ding, M.: 2024, Developing an automated detection, tracking and analysis method for solar filaments observed by CHASE via machine learning. arXiv preprint arXiv.

Zhou, X., Koltun, V., Krähenbühl, P.: 2020, Tracking objects as points. In: *European Conference on Computer Vision*, Springer, Berlin, 474.