Convective Cloud Detection From Himawari-8 Advanced Himawari Imager Data Using a Dual-Branch Deformable Convolutional Network

Renlong Hang, *Member, IEEE*, Jingquan Wang, Lingling Ge, Chunxiang Shi, and Jianfen Wei

***Abstract*—Himawari-8 satellite, equipped with an advanced Himawari imager (AHI), has been widely employed for cloud detection tasks due to its high-spatiotemporal resolution. In this article, we propose a deep learning model named dual-branch deformable convolutional network (DBDCN) to detect convective cloud from AHI data. Specifically, we first choose some infrared channels from AHI to compute their difference. These brightness temperature difference (BTD) data are sensitive to cloud detection, thus providing complementary information for AHI. Then, both AHI and BTD are used as inputs of DBDCN, which adopts a dual-branch structure to effectively combine them together. Considering that convective clouds often exist in variable shapes, we use deformable convolutions to extract convective clouds' features in each branch. In addition, in order to fuse the complementary features from both branches, we propose a crossbranch fusion block. To evaluate the performance of our proposed DBDCN, we collect a convective cloud dataset, covering most of regions in southern China (i.e., 97°E to 122°E, 20°N to 35°N). Experimental results show that DBDCN is able to achieve better detection performance than the widely used threshold methods and existing convolutional neural networks in most cases.**

***Index Terms*—Brightness temperature difference (BTD), convective cloud detection, crossbranch fusion, deformable convolution.**

I. Introduction

CONVECTIVE clouds, a distinct cloud type, significantly influence the global hydrological cycle and the radiative energy balance at local and global scales [1], [2]. They are frequently associated with severe weather conditions, such as heavy rain and hail, posing a substantial threat to both human life and property [3], [4], [5], [6]. Meteorological satellites, such as Himawari-8, have been widely applied to monitoring convective clouds, because they are capable of observing large-scale and continuous cloud information.

In order to effectively detect convective clouds from satellite data, a lot of algorithms have been proposed in the past decades. Based on the observation that convective clouds often have a relatively low cloud top temperature, brightness temperature (BT) threshold method using infrared channel is a popularly adopted algorithm [7], [8], [9], [10]. For example, Maddox [11] chose the infrared channel at 10.7 μm to identify regions characterized by strong convective clouds. Fang et al. [12] designed a threshold in the water vapor channel to distinguish convective clouds from other clouds. In addition to BT data, BT difference (BTD) data between some chosen channels are also widely used, because they are sensitive to some kinds of clouds, whereas BT are not [13]. In [14], the infrared channel and the water vapor channel were chosen to compute the difference data, which can describe water droplets or ice particles differences. Bai et al. [15] introduced a threshold method utilizing infrared channel and infrared BTD for the detection of convective clouds. However, the above threshold based detection methods heavily depend on the experts' prior knowledge to select suitable channels and thresholds, which often costs much time. More importantly, the selected thresholds do not always work, because the convective clouds change dramatically from their birth to dissipation [16], [17].

In contrast with threshold methods, deep learning, as a data-driven method, can automatically learn effective features for cloud detection tasks from input data, thus, attracting more and more attention in recent years [18], [19], [20]. For example, Kim et al. [21] introduced the use of convolutional neural network (CNN) to detect cloud information by combining data from infrared and visible channels. Building upon Kim's concept, Lee et al. [22] suggested the utilization of 2-D CNN and 3-D CNN to integrate visible and infrared channels, resulting in improved detection results. In addition, Shao et al. [23] proposed a multiscale feature extraction network, which integrates low-level spatial information and high-level semantic features to achieve pixel-level thin cloud detection. In summary, these CNN-based models often adopt visible or infrared data as inputs and use

Manuscript received 19 December 2023; revised 6 March 2024; accepted 13 March 2024. Date of publication 22 March 2024; date of current version 4 April 2024. This work was supported in part by the Natural Science Foundation of China under Grant U21B2049, in part by the Natural Science Foundation of Jiangsu Province, China, under Grant 23KJB170024, in part by the Advance Research on Civil Space Technology During China's 14th Five-Year Plan under Grant D040405, and in part by the Wuxi University Research Start-up Fund for Introduced Talents under Grant 2023r038. *(Corresponding author: Jianfen Wei.)*

Renlong Hang and Jingquan Wang are with the Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: renlong_hang@163.com; jingquanjq@163.com).

Lingling Ge and Chunxiang Shi are with the National Meteorological Information Center, Beijing 100081, China (e-mail: gell@cma.gov.cn; shicx@cma.gov.cn).

Jianfen Wei is with the Key Laboratory of Ecosystem Carbon Source and Sink, China Meteorological Administration (ECSS-CMA), Wuxi University, Wuxi 214105, China (e-mail: weijianfen0425@163.com).

Digital Object Identifier 10.1109/JSTARS.2024.3380631

convolutional operators to extract clouds' features. However, visible and infrared data are not enough to distinguish convective clouds from other clouds having similar cloud top heights. In addition, the convolutional kernels in traditional convolutions have fixed shapes, making them difficult to learn the diverse shape information in convective clouds.

In order to address the aforementioned challenges, we propose a dual branch deformable convolutional network (DBDCN) for convective cloud detection. It takes not only visible and infrared data but also BTD data as inputs. To effectively extract features from these two kinds of data, a two branch network is adopted. In each branch, deformable convolutions are used to capture the variable shape information of convective clouds. Based on the extracted features, a crossbranch fusion block (CFB) is designed to fuse the complementary information between different branches. To sum up, the primary contributions of this article are as follows.

1) To effectively detect convective clouds, we compute some BTD data sensitive to cloud analysis as complementary inputs of visible and infrared data, and propose a dual branch convolutional network for fusing these data sufficiently.
2) Considering that convective clouds often have variable shapes, deformable convolutions are adopted to extract convective clouds' features in each branch. Based on the extracted features, a CFB module is designed to combine them in different branches.
3) In order to evaluate the performance of different models, a convective cloud dataset based on Himawari-8 satellite is constructed. The quantitative and qualitative results show that our proposed model can achieve better results than several state-of-the-art models in most cases.

II. Data

A. Himawari-8 Advanced Himawari Imager (AHI) Data

On 7 October, 2014, the Japan Meteorological Agency (JMA) initiated the operation of Himawari-8, a state-of-the-art geostationary satellite. Positioned at 140.7°E, this satellite encompasses the east Asian monsoon region, spanning longitudes from 80°E to 160°W and latitudes from 60°N to 60°S. Himawari-8 is equipped with the AHI instrument, which is capable of capturing data from visible to infrared channels, including three visible channels, three near-infrared channels, and ten infrared channels [24], [25]. Different from other geostationary satellites, AHI has a spatial resolution 0.5 to 2 km and a relatively high temporal resolution of 2.5 to 10 min [26]. Such high-spatiotemporal resolution makes it possible to accurately detect convective clouds [27], [28], [29]. For the detailed information about AHI, please refer to the study in [30]. In this article, all channels of AHI are used for the task of convective cloud detection.

B. Study Area

China experiences the influence of the summer monsoon, with the majority of annual precipitation concentrated in the summer and autumn months, specifically from May to September. During this period, precipitation in the southern China generally surpasses that in the northern regions. In addition, according to the analysis of the China Climate Bulletin,[1,2] the annual precipitation in 2020 and 2021 is significantly higher than the average annual precipitation. Therefore, we choose the southern China, spanning from 97°E to 122°E and 20°N to 35°N, as our study area, and the period from May to September in both 2020 and 2021 as our study time. Fig. 1 illustrates the distribution of our study area with a red box. In the experiments, we use the data in the year 2020 to train our model and the data in the year 2021 to test the performance of our model.

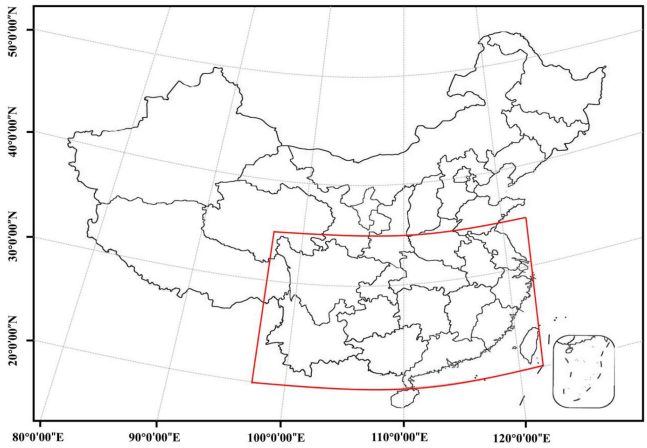

Fig. 1. Illustration of the study area we choose, covering most of the southern China from 97°E to 122°E longitude and 20°N to 35°N latitude.

C. Convective Clouds Labeling

The Meteorological Satellite Center of the JMA has introduced a product called high-resolution cloud analytical information (HCAI) [31], which consists of five elements: cloud mask, snow-ice mask, cloud type, cloud top height, and quality control information. For the cloud type product, three infrared channels (i.e., 6.2, 7.3, and 10.4 μm) are used to distinguish different kinds of clouds, including cumulonimbus clouds, a typical convective cloud. It is reasonable to choose cloud type product as an initial convective clouds label for our study. Due to the strict threshold selection during the generation of this product, there often exist many holes in a detected cumulonimbus cloud. Fig. 2(c) shows an instance of such situation. In addition, the cloud type product may ignore other kinds of convective clouds. In order to address these issues, we take advantage of visible and infrared channels in AHI to manually fine-tune the labels via visual interpretation. As shown in Fig. 2(e), more numbers of homogeneous convective clouds are labeled. The small-scaled convective clouds, which are easily overlooked by the cloud type product, are also correctly labeled.

[1][Online]. Available: https://www.cma.gov.cn/zfxxgk/gknr/qxbg/202104/P020230712370988699486.pdf

[2][Online]. Available: https://www.cma.gov.cn/zfxxgk/gknr/qxbg/202203/P020230712370155615526.pdf

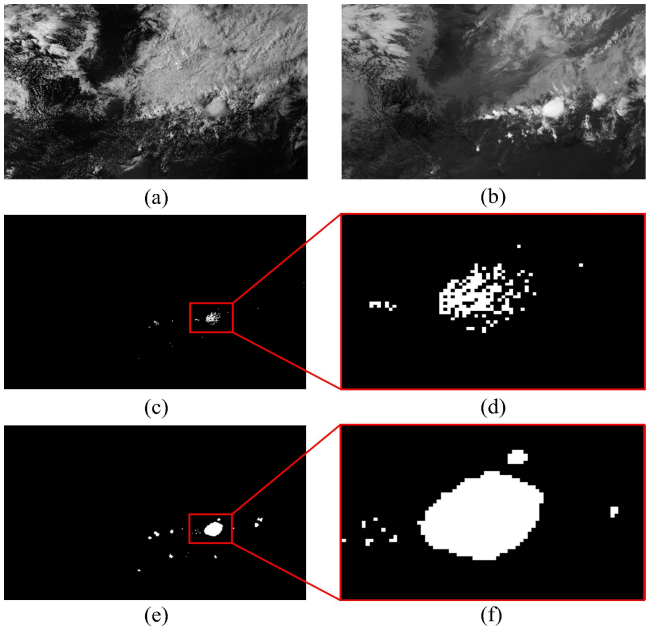

Fig. 2. Illustration of convective cloud labels. (a) Visible channel data. (b) Infrared channel data. (c) and (d) HCAI detected convective clouds and their locally zoomed results. (e) and (f) Manually fine-tuned results and their locally zoomed results.

D. BTD Data

Through the analysis of AHI data and its products [32], we choose infrared channels at 3.9, 6.2, 7.3, 8.6, 10.4, 11.2, and 12.4 μm to construct BTD data. For simplicity, these channels are abbreviated as BT3.9, BT6.2, BT7.3, BT8.6, BT10.4, BT11.2, and BT12.4, respectively.

The first BTD data is derived by BT7.3–BT11.2. It can help detect the distribution of water vapor content in the cloud. The larger the difference is, the higher the water vapor content is, which indicates that the cloud is more likely to be a convective cloud [33].

The second BTD data is BT11.2–BT3.9. BT11.2 can be used to observe high clouds, whereas BT3.9 is mainly adopted to observe low clouds. Therefore, their difference is beneficial to detect the cloud thickness. When the difference is larger, the cloud is thicker [33].

The third BTD data is BT11.2–BT12.4, which can detect the height of cloud top. The larger difference value represents that the cloud is a convective cloud with a higher probability [33].

The fourth BTD data is BT8.6–BT11.2. This difference value can help detect ice clouds. When such value is larger, the cloud is an ice cloud with a higher probability [34].

The fifth BTD data is BT6.2–BT10.4. BT6.2 and BT10.4 are employed to observe cloud bottom and cloud top BTs, respectively. Their difference indicates the vertical structure information of clouds, and the large difference values represent that the clouds are highly developed [35].

The last BTD data is BT10.4–BT12.4, which reflects the water vapor content in the clouds. The cloud is more likely to be a convective cloud when the difference value is larger [36].

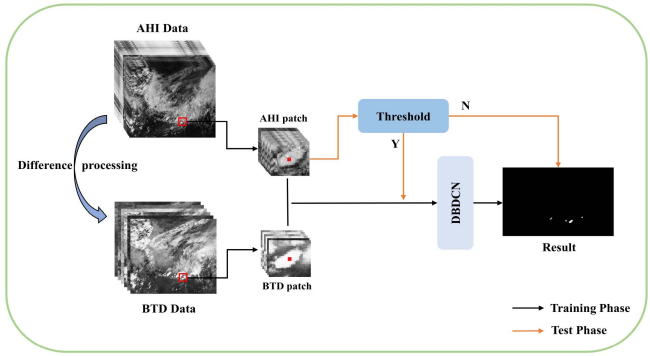

Fig. 3. Flowchart of our proposed convective cloud detection method.

III. Methodology

Fig. 3 shows the framework of our proposed convective cloud detection method. Assume the given AHI data is $X \in R^{H\times W\times C}$, where H, W, and C represent height, width, and the number of channels in AHI, respectively. We first compute the BTD data $Y \in R^{H\times W\times C'}$ according to Section II-D, where C' denotes the numbers of BTD data. Then, for each pixel in X, we crop a patch $x \in R^{p\times p\times C}$ centered at it, where p is the patch size. Meanwhile, at the same spatial position, we crop a patch $y \in R^{p\times p\times C'}$ from Y. In our study, we empirically set p to 7. After that, both x and y are fed into DBDCN to identify whether the input pixel is a convective cloud. Such pixelwise identification method needs to implement DBDCN $H \times W$ times, which costs much time. To solve this issue, we adopt a simple threshold method to preprocess each pixel. Specifically, we set a threshold in the 11.2 μm channel and compare each pixel's value at the same channel with it. To avoid missing the developing convective cloud, the threshold is empirically set to 225 K. If the BT value of a given pixel in the 11.2 μm channel is larger than the predefined threshold, it is directly identified as nonconvective cloud. Otherwise, the pixel needs to be fed into DBDCN to make an identification.

A. DBDCN Architecture

As discussed above, DBDCN is the most important part of our proposed method. As shown in Fig. 4(a), DBDCN is comprised of two branches, and they are employed to process x and y, respectively. In each branch, a convolutional layer whose kernel size is 5×5 is used to extract an initial feature from the input data. Considering that the shape of different convective clouds often vary, a deformable convolution layer with kernel size 3×3 is applied to capture various shape information. After that, the extracted features $F_1 \in R^{p/2\times p/2\times 2C}$ and $F_1' \in R^{p/2\times p/2\times 2C}$ in both branches are fed into CFB to combine the complementary information between them. Such complementary information is used to further enhance the discriminative ability of F_1 and F_1' via a residual block, which consists of two convolutional layers with kernel size 3×3. Subsequently, the aforementioned process with a deformable convolution, CFB and a residual block are repeatedly implemented twice to extract high-level features F_4 and F_4'. Finally, F_4 and F_4' are concatenated together along

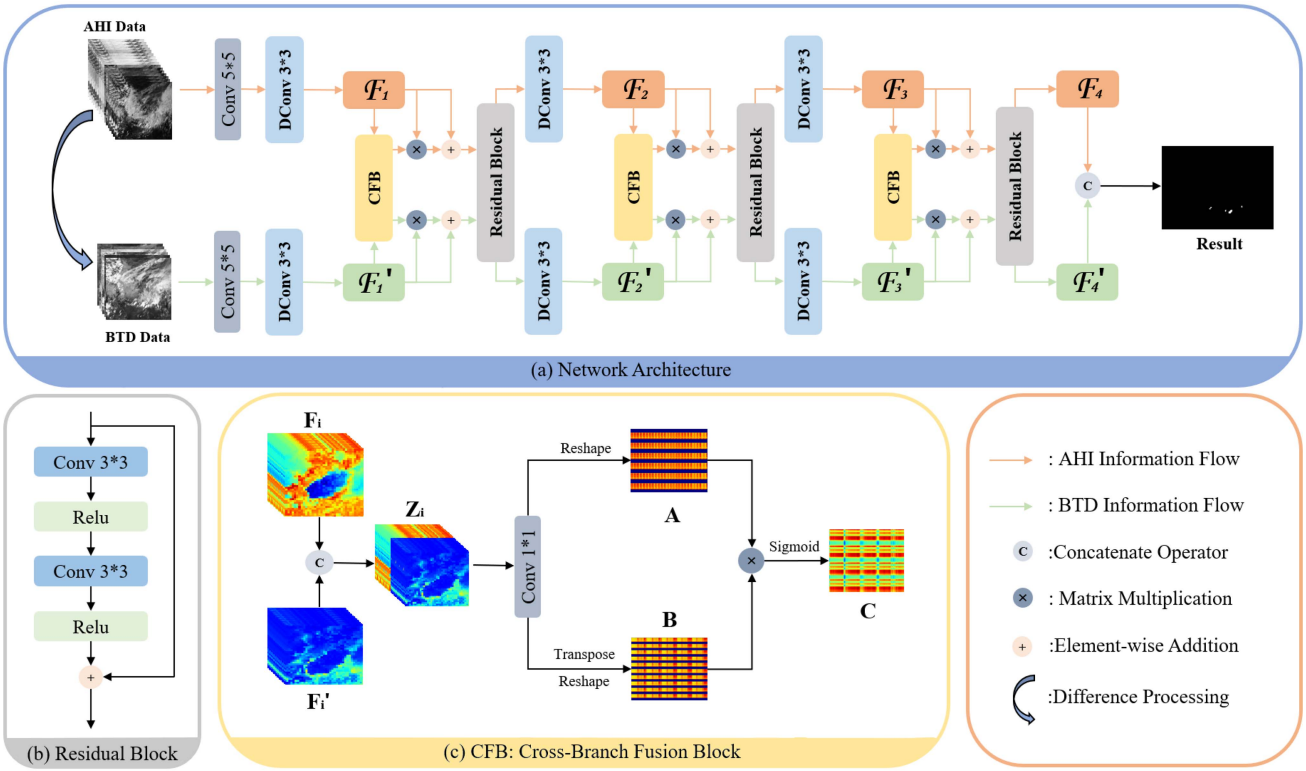

Fig. 4. Detailed architecture of our proposed DBDCN, which mainly consists of deformable convolutions, CFBs and residual blocks. In each branch, deformable convolutions are used to extract features. Such features in both branches are fused together via CFBs. Following deformable convolutions and CFBs, residual blocks are employed to fine-tune the features in each branch.

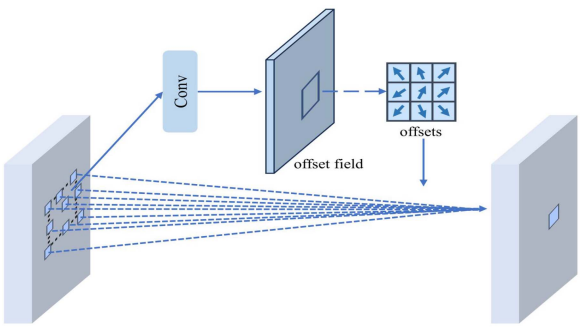

Fig. 5. Detailed architecture of deformable convolution.

the channel dimension, and fed into a fully-connected layer to derive a cloud detection result.

B. Deformable Convolution

Traditional convolution operators have fixed shape kernels, which are difficult to learn the diverse shape information of convective clouds. In contrast, deformable convolution operators can adaptively adjust the shapes of convolutional kernels, making them more suitable to extract features from convective clouds [37]. As shown in Fig. 5, deformable convolution adopts a single-layer convolution to generate a learnable offset, which is used to adjust the position of the kernel at the sampling point on the input. We use the learned offset to calculate the actual position of the sampling point. Assuming that a sampling point on the input feature map is (e, f), and the corresponding offset is $(\Delta e, \Delta f)$, then the new sampling point position is $(e + \Delta e, f + \Delta f)$. Thus, sampling points of the convolution kernel can be adaptively adjusted according to the structure of convective clouds. Fig. 6 shows an instance to demonstrate the difference between the traditional convolution and the deformable convolution. It is obvious to observe that the deformable convolution is able to better fit the shape of convective clouds.

C. Structure of CFB

Fig. 4(c) shows the flowchart of our proposed CFB. For a given pair of features F_i and F'_i, $i \in \{1, 2, 3\}$ from the two branches of DBDCN, we first concatenate them together along the channel dimension to combine the complementary information between them. Since BTD data is derived from some chosen channels of AHI, there may exist some similarity information between F_i and F'_i. Therefore, we use a convolutional layer whose kernel size is 1×1 to reduce the redundancy between different channels. The above process can be formulated as

$$Z_i = \mathrm{Conv}([F_i, F'_i]) \tag{1}$$

where [.,.] represents the concatenation operation, Conv is the 1×1 convolutional operator, and $Z_i \in R^{h \times w \times c}$ is the fusion result.

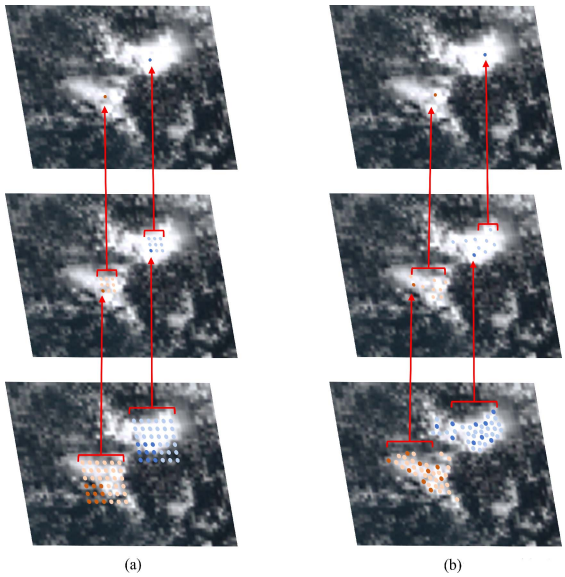

Fig. 6. Comparison between the traditional convolution and the deformable convolution. (a) Traditional convolution. (b) Deformable convolution.

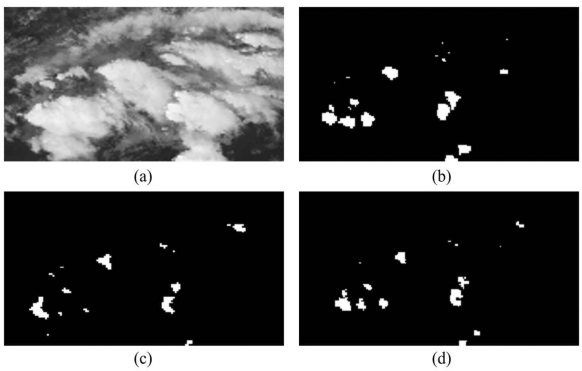

Fig. 7. Visualization of the detection results with and without using BTD data. (a) Infrared data in the 11.2 μm channel. (b) Ground-truth map. (c) Detection map using only AHI data. (d) Detection map using both AHI and BTD data.

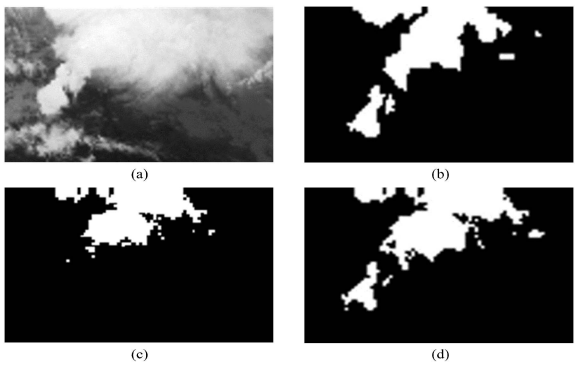

Fig. 8. Visualization of the detection results achieved by different models. (a) Infrared data in the 11.2 μm channel. (b) Ground-truth map. (c) Detection map achieved by the Baseline model. (d) Detection map achieved by the Baseline+DConv model.

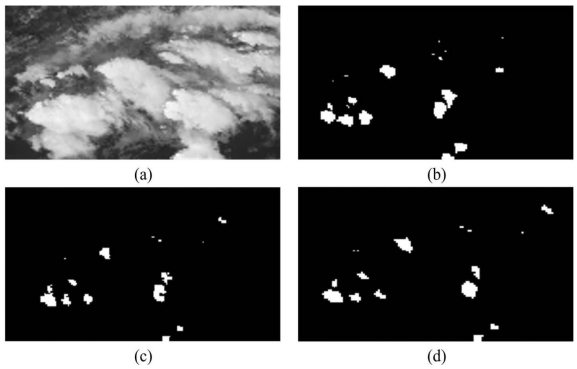

Fig. 9. Visualization of the detection results achieved by different models. (a) Infrared data in the 11.2 μm channel. (b) Ground-truth map. (c) Detection map achieved by the Baseline model. (d) Detection map achieved by the Baseline+CFB model.

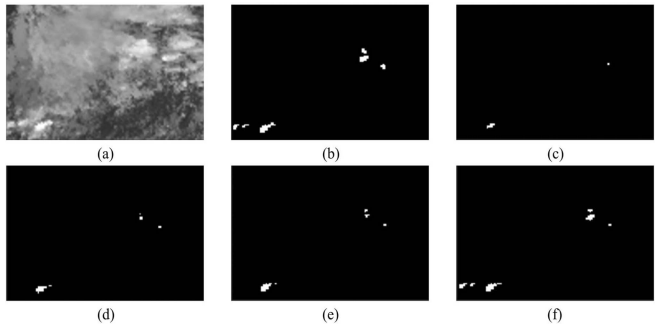

Fig. 10. Visualization of the detection results achieved by different models. (a) Infrared data in the 11.2 μm channel. (b) Ground-truth map. (c) Detection map achieved by the Baseline model. (d) Detection map achieved by the Baseline+DConv model. (e) Detection map achieved by the Baseline+CFB model. (f) Detection map achieved by the DBDCN model.

After acquiring Z_i, we attempt using such complementary information to enhance F_i and F_i' simultaneously. Specifically, we reshape Z_i into a matrix $A \in R^{n \times c}$ and transpose A into another matrix $B \in R^{c \times n}$, where n represents $h \times w$. Then, the relationship between different channels in Z_i can be modeled as

$$C = \sigma(BA) \tag{2}$$

where σ represents the *sigmoid* activation function, and each element in $C \in R^{c \times c}$ denotes the relationship between two corresponding channels. Based on C, F_i, and F_i' are enhanced by $F_i C$ and $F_i' C$, respectively.

IV. Experiments

In this section, we perform a comprehensive evaluation of our proposed model using AHI data. Specifically, we begin by providing a concise overview of the implementation details, followed by the execution of ablation experiments on several critical modules in DBDCN. Lastly, we comprehensively assess the model's performance through both qualitative and quantitative analyses.

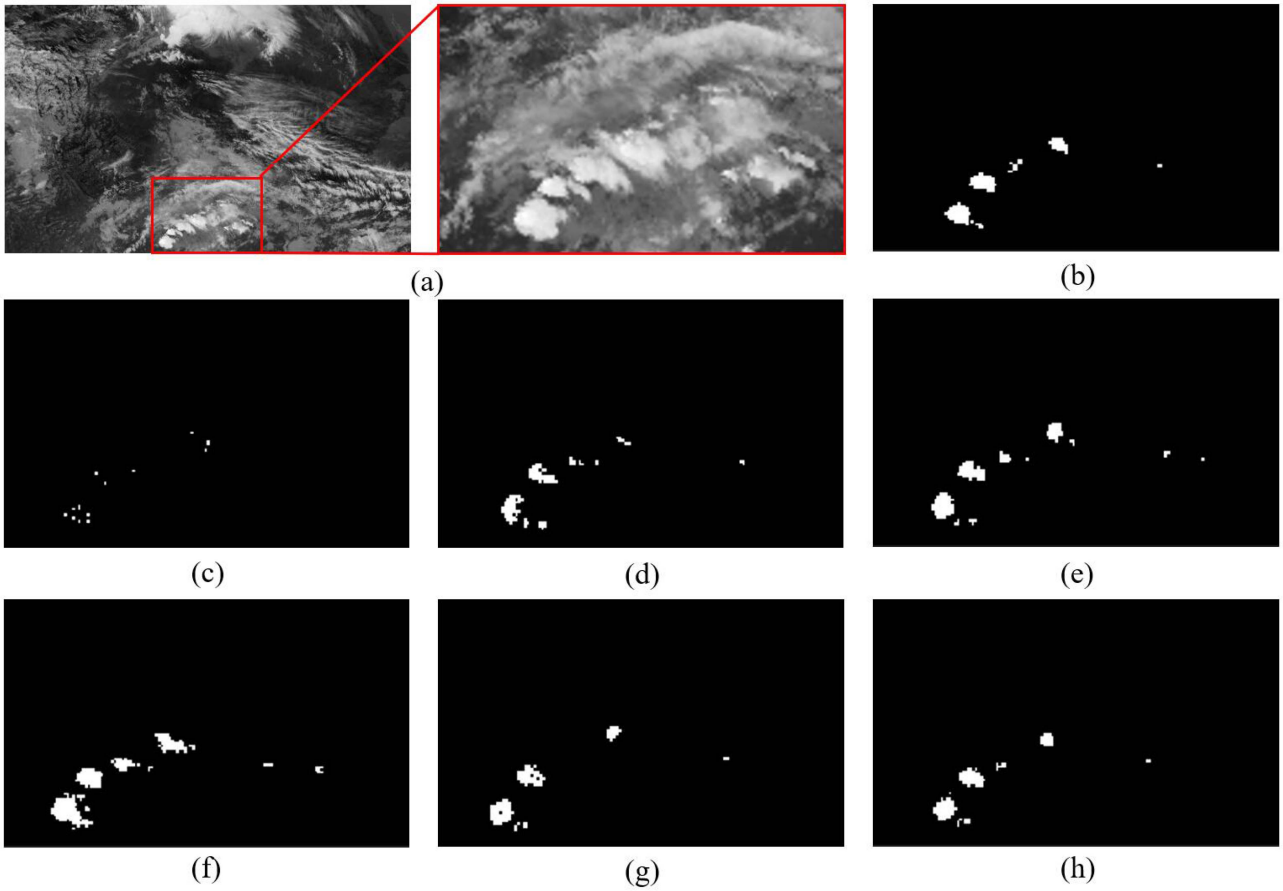

Fig. 11. Visualization of the detection results achieved by different models. (a) Infrared data in the 11.2 μm channel. (b) Ground-truth map. (c)–(h) Detection maps achieved by HCAI, DSFC, Convection-UNet, CDNet, DABNet, and DBDCN, respectively.

A. Implementation Details

1) Data Splitting: As discussed in Section II-B, we collect the data from May to September in both 2020 and 2021 to construct experiments. Among these data, the first 90% data (i.e., 326 samples) in 2020 are used as a training set, whereas the last 10% data (i.e., 36 samples) in 2020 are used as a validation set. All of the data in 2021 (i.e., 210 samples) are used as a test set. The spatial size of each sample is 301×501. To train our proposed model effectively, each sample in the training and validation sets is cropped into patches, resulting in 167 000 and 47 000 patches in the training set and the validation set, respectively.

2) Evaluation Metrics: To assess the performance of different methods in the task of convective cloud detection, we adopt four commonly used quantitative indicators : probability of detection (POD), missed alarm rate (MAR), false alarm rate (FAR), and critical success index (CSI) [38]. These indicators are calculated using the following formulas:

$$\text{POD} = \frac{\text{TP}}{TP + FN} \tag{3}$$

$$MAR = 1 - \text{POD} \tag{4}$$

$$FAR = \frac{\text{FP}}{\text{TP} + \text{FP}} \tag{5}$$

$$CSI = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{6}$$

where TP and TN represent the count of correctly detected convective clouds and nonconvective clouds, respectively, while FP and FN denote the count of convective clouds and nonconvective clouds incorrectly detected, respectively. It is important to highlight that POD and CSI serve as measures of detection accuracy, with higher values indicating better results, whereas MAR and FAR represent error detection, with smaller values signifying better results.

3) Comparison Models: In this study, we compare our proposed model with several representative CNN-based cloud detection methods: DSFC [39], Convection-UNet [40], CDNet [41], and DABNet [42]. Furthermore, we also make a comparison with the cloud classification product HCAI [31]. DSFC is a dual-branch network, which uses 6-layer 1×1 convolutions and 2-layer 3×3 convolutions to extract spectral and spatial features, respectively. Convection-UNet adopts an encoder–decoder framework. In the encoder, 10-layer 3×3 convolutions are used to extract features. In the decoder, 8-layer 3×3 deconvolutions are used to recover the resolution of the feature.

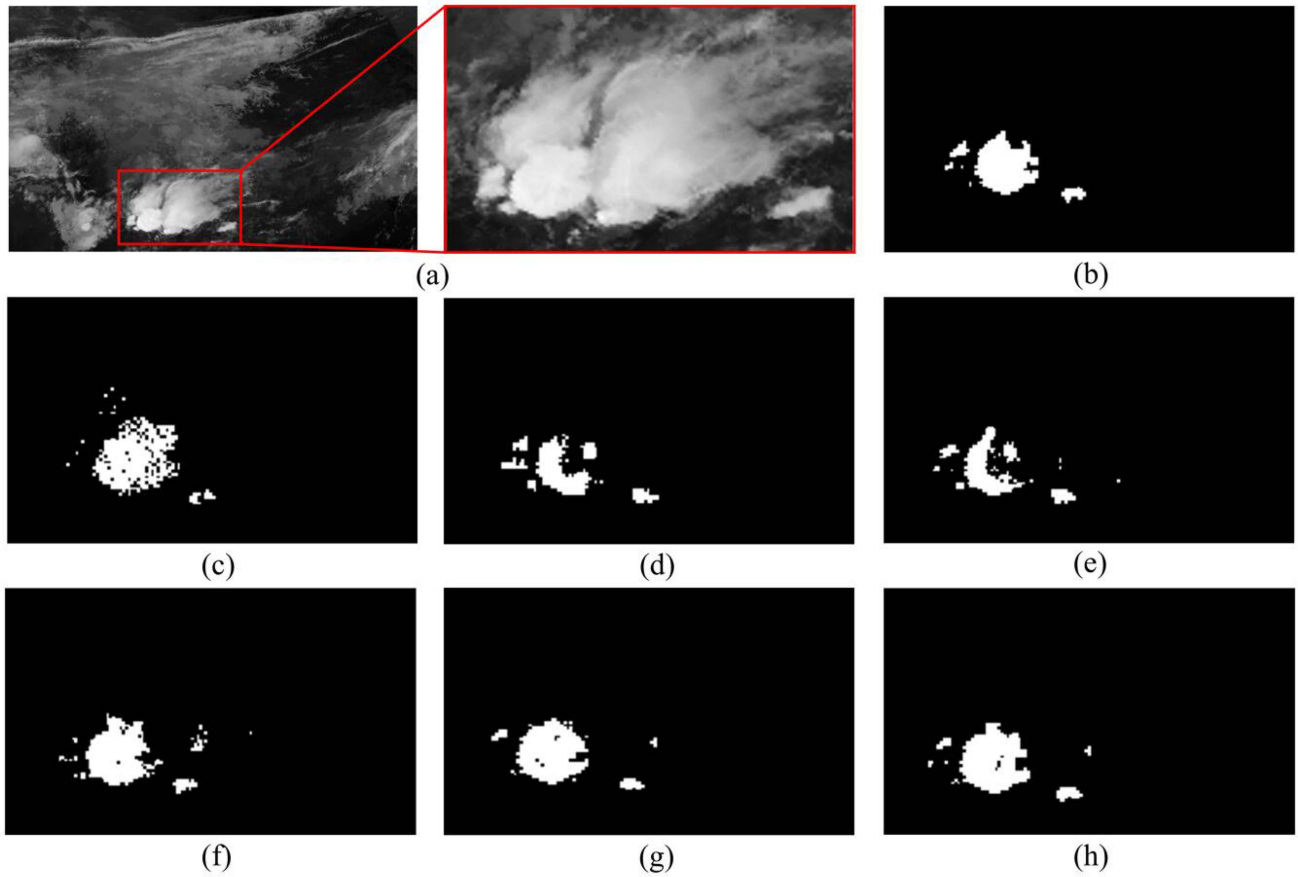

Fig. 12. Visualization of the detection results achieved by different models. (a) Infrared data in the 11.2 μm channel. (b) Ground-truth map. (c)–(h) Detection maps achieved by HCAI, DSFC, Convection-UNet, CDNet, DABNet, and DBDCN, respectively.

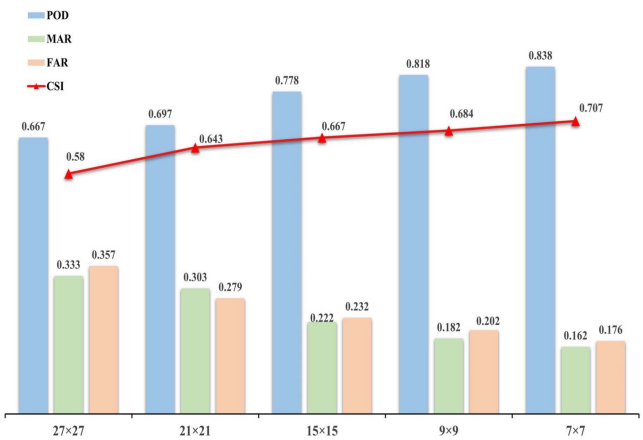

Fig. 13. Effects of different patch sizes on the detection performance.

CDNet uses ResNet-50 as backbone, and then 3×3 and 1×1 convolutions are used to build a multiscale feature extraction module to fuse features at different levels. DABNet uses 8-layer 3×3 dilated convolutions to extract features.

4) Experimental Setting: All models are trained using the training set and their performance are evaluated on the test set. In order to optimize the parameters of each model, we use an adaptive moment estimation (Adam) algorithm as the optimizer and crossentropy loss as the loss function. The learning rate is set to 1e-4, and each model is trained for a total of 50 epochs. All the experiments are implemented with a NVIDIA RTX 3090 GPU under the PyTorch framework.

B. Ablation Study

In our proposed DBDCN, there are three critical components, including the use of BTD data, deformable convolution (DConv) and CFB. Here, we conduct ablation studies to show the effectiveness of them.

1) Effect of BTD Data: To assess the effects of BTD data on the detection performance, we conduct experiments with two different networks. The first one is taking the AHI data as input, whereas the second one has both AHI and BTD data as inputs. It is worth noting that both networks do not use DConv layers and CFB modules to simply experiment. Table I quantitatively compare their detection performance. It can be observed that using AHI and BTD together achieves higher performance, in terms of POD, MAR, FAR, and CSI, than the individual use of AHI. In addition, Fig. 7 qualitatively shows some detection maps obtained by using or not using BTD data. In comparison with the ground-truth map in Fig. 7(b), the network without using BTD data misidentifies some regions in the top right corner as

TABLE I
EFFECTS OF BTD DATA ON CONVECTIVE CLOUD DETECTION

Input data	POD ↑	MAR ↓	FAR ↓	CSI ↑
AHI	0.709	0.291	0.358	0.500
+ BTD	**0.721**	**0.279**	**0.247**	**0.571**

The best accuracies are shown in bold values.

TABLE II
EFFECTS OF DCONV AND CFB ON CONVECTIVE CLOUD DETECTION

Models	POD ↑	MAR ↓	FAR ↓	CSI ↑
Baseline	0.721	0.279	0.247	0.571
+ DConv	0.755	0.245	0.213	0.652
+ CFB	0.764	0.236	0.221	0.661
DBDCN	**0.838**	**0.162**	**0.176**	**0.707**

The best accuracies are shown in bold values.

convective clouds and ignores some convective cloud regions in the left bottom of Fig. 7(c). Differently, when BTD data is added, the network, shown in Fig. 7(d), is able to obtain more consistent results in these regions with the ground truth. Therefore, we can conclude that using BTD data is beneficial to the task of convective cloud detection.

2) Effect of DConv: To evaluate the effects of DConv on the detection performance, we first construct a " Baseline" model, which removes CFB modules and DConv layers from DBDCN. Then, we add DConv to the Baseline model. Table II and Fig. 8 compare the detection performance between Baseline and Baseline+DConv. As demonstrated in the first two rows of Table II, Baseline+DConv increases the POD and CSI values by about 5%, and decreases the MAR and FAR values by more than 10%, when compared with Baseline. Also, in comparison with the ground-truth map in Fig. 8(b), Baseline+DConv obtains a more complete shape of convective clouds than Baseline. Both of the aforementioned quantitative and qualitative results certify the effectiveness of the use of DConv.

3) Effect of CFB: Similar to the above section, we add CFB module to the Baseline model to test the effects of it on the detection performance. As reported in the first and the third rows of Table II, Baseline+CFB is able to improve the detection performance significantly in terms of all four indicators when compared with Baseline. The detection maps shown in Fig. 9 also demonstrate that Baseline+CFB obtains more consistent maps than Baseline. All of these observations validate the effectiveness of using CFB modules.

4) Effect of DConv and CFB: In addition to evaluate the individual effects of using DConv and CFB, we also evaluate the effects of their combinations. To this end, we compare DBDCN with other models. The comparison results are presented in Table II and Fig. 10. In comparison with both Baseline+DConv and Baseline+CFB, DBDCN can improve the performance in terms of POD, MAR, FAR, and CSI values by about 10%, 30%, 17%, and 7%, respectively. As shown in Fig. 10(c), Baseline overlooks the small-scale convective clouds in the left bottom and the top right regions. Both Baseline+DConv and Baseline+CFB models alleviate this issue partly, but the convective clouds are still not perfectly detected, especially in the left bottom region. Different from them, DBDCN detects almost all convective clouds. All of these results indicate the importance of using both DConv and CFB.

C. Comparison With State-of-The-Art Models

1) Quantitative Comparison: Table III reports the performance achieved by different models, where bold fonts represent the best performance in each indicator. Several conclusions can be observed from this table. First of all, due to the strict threshold selection, many convective clouds are not detected by HCAI, making it achieve the worst detection performance in comparison with other models. Second, for the four deep learning based detection models, CDNet obtains the best POD and MAR values, whereas DABNet achieves the best FAR and CSI values. Compared with other deep learning methods, CDNet designed a pyramid structure to fuse shallow and deep convolution features in order to prevent feature loss. A residual structure is also added subsequently to each upsampling to help recover lost features. In contrast with CDNet, DABNet pays more attention to the detection accuracy of the cloud boundary. Besides introducing deformable convolution, DABNet also designs a new loss function to solve the boundary ambiguity problem. Last but not the least, our proposed DBDCN can yield the best performance in three of four indicators. Specifically, in comparison with CDNet, DBDCN improves the POD, MAR, FAR, and CSI values by about 8.7%, 29.3%, 12.9%, and 2.9%, respectively. In comparison with DABNet, although DBDCN achieves a slightly worse FAR value, it improves the POD, MAR, and CSI values by 9.3%, 30.5%, and 1.4%, respectively. These conclusions can validate the effectiveness of our proposed DBDCN.

2) Qualitative Comparison: Figs. 11 and 12 visualize the detection maps achieved by different models. Since convective clouds only cover a small portion of the whole satellite image, we crop the convective cloud region and enlarge it to make a more clear comparison. In these detection maps, the white and black colors represent the detected convective clouds and nonconvective clouds, respectively. In comparison with the ground-truth map, it can be observed that HCAI ignores most of convective clouds in Fig. 11(c) and many holes exist in the detected clouds of Fig. 12(c). DSFC obtains better detection maps than HCAI, but it only detects a part of the whole clouds. Due to the exploit of multiscale convolution features, Convection-UNet and CDNet detect clouds in different scales, but the cloud shapes are not good enough. Differently, as shown in Figs. 11(g) and 12(g), DABNet can capture better cloud shapes than Convection-UNet and CDNet, but some small-scale clouds are not detected. Compared with all other models, our proposed DBDCN can detect different-scale clouds and their shapes are well preserved.

TABLE III
PERFORMANCE COMPARSION BETWEEN DIFFERENT MODELS

Method	Ref.	Year	POD ↑	MAR ↓	FAR ↓	CSI ↑
HCAI	[31]	2015	0.338	0.662	0.383	0.238
DSFC	[39]	2019	0.595	0.405	0.292	0.460
Convection-UNet	[40]	2023	0.679	0.321	0.314	0.492
CDNet	[41]	2019	0.771	0.229	0.202	0.687
DABNet	[42]	2022	0.767	0.233	**0.173**	0.697
DBDCN	-	-	**0.838**	**0.162**	0.176	**0.707**

The best accuracies are shown in bold values.

TABLE IV
EFFECTS OF DIFFERENT THRESHOLD VALUES ON THE DETECTION PERFORMANCE

Threshold	POD ↑	MAR ↓	FAR ↓	CSI ↑	Time(s)
230K	0.764	0.236	0.204	0.684	130
225K	**0.838**	**0.162**	**0.176**	**0.707**	100
220K	0.749	0.251	0.210	0.676	**75**

The best accuracies are shown in bold values.

D. Discussion

1) Patch Size: The input of our proposed DBDCN is cropped patches centered at each pixel. Considering that convective clouds often have relatively small scales, the cropped patch sizes cannot be set too large. To choose a suitable size, we conduct experiments with five different sizes, including 27×27, 21×21, 15×15, 9×9, and 7×7, and test their effects on the detection performance of our proposed DBDCN. As shown in Fig. 13, when the patch size is decreased from 27×27 to 7×7, MAR and FAR values will decrease, whereas POD and CSI values will increase, so 7×7 is a better choice for DBDCN.

2) Threshold Selection: In order to reduce the computation time of our proposed method, we adopt a threshold method to preprocess each pixel before feding into DBDCN. When a larger threshold is applied, only a smaller number of pixels are filtered as non-convective clouds, so the reduced computation time is not much. On the contrary, a smaller threshold will filter larger numbers of pixels, but the convective clouds will be filtered also in a high probability. To make a good balance between computation time and detection performance, we conduct some experiments to find a suitable threshold. As demonstrated in Table IV, the threshold setting to 225 K is a better choice than the others.

V. CONCLUSION

In this article, a DBDCN was proposed to detect convective clouds from Himawari-8 satellite. The inputs of DBDCN were AHI data and BTD data derived from some chosen channels of AHI. Such two kinds of input data were processed by the two branches of DBDCN, respectively. In each branch, considering the variable shapes of convective clouds, deformable convolutions were adopted to extract their features, which can adaptively adjust convolutional kernels' shapes. These extracted features were then enhanced by a CFB, which can learn the complementary information from different branches. In order to test the performance of DBDCN, a convective cloud dataset was collected and several comparison experiments were conducted on it. In addition, some ablation study experiments were implemented to evaluate the effects of different components and hyperparameters in DBDCN.

REFERENCES

[1] Z. Feng et al., "Spatiotemporal characteristics and large-scale environments of mesoscale convective systems east of the rocky mountains," *J. Climate*, vol. 32, no. 21, pp. 7303–7328, 2019.
[2] L. M. Carvalho, C. Jones, and B. Liebmann, "Extreme precipitation events in southeastern south America and large-scale convective patterns in the south atlantic convergence zone," *J. Climate*, vol. 15, no. 17, pp. 2377–2394, 2002.
[3] D. J. Boccippio, K. L. Cummins, H. J. Christian, and S. J. Goodman, "Combined satellite-and surface-based estimation of the intracloud–cloud-to-ground lightning ratio over the continental United States," *Monthly Weather Rev.*, vol. 129, no. 1, pp. 108–122, 2001.
[4] L. C. Ivers and E. T. Ryan, "Infectious diseases of severe weather-related and flood-related natural disasters," *Curr. Opin. Infect. Dis.*, vol. 19, no. 5, pp. 408–414, 2006.
[5] M. Baldauf, A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, "Operational convective-scale numerical weather prediction with the cosmo model: Description and sensitivities," *Monthly Weather Rev.*, vol. 139, no. 12, pp. 3887–3905, 2011.
[6] Q. Long, Q. Chen, K. Gui, and Y. Zhang, "A case study of a heavy rain over the southeastern Tibetan plateau," *Atmosphere*, vol. 7, no. 9, 2016, Art. no. 118.
[7] A. M. Haberlie and W. S. Ashley, "A radar-based climatology of mesoscale convective systems in the United States," *J. Climate*, vol. 32, no. 5, pp. 1591–1606, 2019.
[8] X. Huang et al., "A long-term tropical mesoscale convective systems dataset based on a novel objective automatic tracking algorithm," *Climate Dyn.*, vol. 51, pp. 3145–3159, 2018.
[9] L. Machado, W. Rossow, R. Guedes, and A. Walker, "Life cycle variations of mesoscale convective systems over the americas," *Monthly Weather Rev.*, vol. 126, no. 6, pp. 1630–1654, 1998.
[10] Z. Feng et al., "Structure and evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-permitting simulations over the United States," *J. Adv. Model. Earth Syst.*, vol. 10, no. 7, pp. 1470–1494, 2018.
[11] R. A. Maddox, "Mesoscale convective complexes," *Bull. Amer. Meteorological Soc.*, vol. 61, no. 11, pp. 1374–1387, 1980.

[12] X. Fang, H. Qiu, Z. Cao, X. Wang, and G. Hong, "Application of AMSU b microwave data to identify strong convective cloud area," *Meteorological Monthly*, vol. 34, no. 3, 2008, Art. no. 9.
[13] X. Zhuge and X. Zou, "Summertime convective initiation nowcasting over southeastern China based on advanced Himawari imager observations," *J. Meteorological Soc. Japan, Ser.II*, vol. 96, no. 4, 337–353, 2018.
[14] M. Setvák, R. M. Rabin, and P. K. Wang, "Contribution of the modis instrument to observations of deep convective storms and stratospheric moisture detection in goes and MSG imagery," *Atmospheric Res.*, vol. 83, no. 2–4, pp. 505–518, 2007.
[15] J. Bai, H. Wang, and Z. Tao, "The recognition and tracking of severe convective cloud from ir images of GMS," *J. Trop. Meteorol.*, no. 2, pp. 192–201, 1997.
[16] Y. Ai, J. Li, W. Shi, T. J. Schmit, C. Cao, and W. Li, "Deep convective cloud characterizations from both broadband imager and hyperspectral infrared sounder measurements," *J. Geophys. Res.: Atmos.*, vol. 122, no. 3, pp. 1700–1712, 2017.
[17] D. W. Martin, R. A. Kohrs, F. R. Mosher, C. M. Medaglia, and C. Adamo, "Over-ocean validation of the global convective diagnostic," *J. Appl. Meteorol. Climatol.*, vol. 47, no. 2, pp. 525–543, 2008.
[18] X. Li, X. Yang, X. Li, S. Lu, Y. Ye, and Y. Ban, "GCDB-UNet: A novel robust cloud detection approach for remote sensing images," *Knowl.-Based Syst.*, vol. 238, 2022, Art. no. 107890.
[19] M. Tian, H. Chen, and G. Liu, "Cloud detection and classification for S-NPP FSR cris data using supervised machine learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 9827–9830.
[20] J. H. Jeppesen, R. H. Jacobsen, F. Inceoglu, and T. S. Toftegaard, "A cloud detection algorithm for satellite imagery based on deep learning," *Remote Sens. Environ.*, vol. 229, pp. 247–259, 2019.
[21] M. Kim, J. Lee, and J. Im, "Deep learning-based monitoring of overshooting cloud tops from geostationary satellite data," *GIScience Remote Sens.*, vol. 55, no. 5, pp. 763–792, 2018.
[22] J. Lee, M. Kim, J. Im, H. Han, and D. Han, "Pre-trained feature aggregated deep learning-based monitoring of overshooting tops using multi-spectral channels of geokompsat-2a advanced meteorological imagery," *GIScience Remote Sens.*, vol. 58, no. 7, pp. 1052–1071, 2021.
[23] Z. Shao, Y. Pan, C. Diao, and J. Cai, "Cloud detection in remote sensing images based on multiscale features-convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 4062–4076, Jun. 2019.
[24] H. Letu et al., "Ice cloud properties from Himawari-8/AHI next-generation geostationary satellite: Capability of the AHI to monitor the dc cloud generation process," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3229–3239, Jun. 2018.
[25] H. Letu et al., "High-resolution retrieval of cloud microphysical properties and surface solar radiation using himawari-8/ahi next-generation geostationary satellite," *Remote Sens. Environ.*, vol. 239, 2020, Art. no. 111583.
[26] A. Okuyama et al., "Preliminary Validation of Himawari-8/AHI navigation and calibration," in *Proc SPIE Earth Observing Syst.*, vol. 9607, 2015, pp. 663–672.
[27] Y. Yang, C. Zhao, Y. Sun, Y. Chi, and H. Fan, "Convective cloud detection and tracking using the new-generation geostationary satellite over south China," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4103912.
[28] Z. Wang, X. Kong, Z. Cui, M. Wu, and T. Liu, "VECNET: A spectral and multi-scale spatial fusion deep network for pixel-level cloud type classification in Himawari-8 imagery," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2021, pp. 4083–4086.
[29] S. Zha, W. Jin, C. He, Z. Chen, G. Si, and Z. Jin, "Detecting of overshooting cloud tops via himawari-8 imagery using dual channel multiscale deep network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1654–1664, 2020.
[30] K. Bessho et al., "An introduction to Himawari-8/9- Japan's new-generation geostationary meteorological satellites," *J. Meteorological Soc. Japan. Ser. II*, vol. 94, no. 2, pp. 151–183, 2016.
[31] T. I. Hiroshi Suzue and K. Mouri, "High-resolution cloud analysis information derived from himawari-8 data," Meteorological Satellite Center, Beijing China, Technical Note no. 61, Mar. 2016.
[32] T. Imai and Y. Ryoshi, "Algorithm theoretical basis for Himawari-8 cloud mask product," *Meteorological Satell. Center Tech. Note*, no. 61, pp. 1–17, Mar. 2016.
[33] N. An, H. Shang, D. Huslert, Q. Hai, and Y. Bao, "Research on night-time cloud identification method based on Himawari image in China," *Equip. Environ. Eng.*, vol. 16, no. 6, 2019, Art. no. 8.
[34] P. V. Hobbs and A. L. Rangno, "Ice particle concentrations in clouds," *J. Atmospheric Sci.*, vol. 42, no. 23, pp. 2523–2549, 1985.
[35] Z. Yiqin, Y. Xiaofeng, and L. Ziwei, "Detection of severe convective cloud over sea surface from geostationary meteorological satellite images based on deep learning," *J. Remote Sens.*, vol. 24, no. 1, pp. 97–106, 2020.
[36] Y. Tingya and W. Shan, "The ghsom network cloud classification model of stationary satellite infrared cloud images at night," *J. Jiangxi Normal Univ. (Natural Sci. Ed.)*, vol. 39, no. 4, pp. 383–388, 2015.
[37] J. Dai et al., "Deformable convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 764–773.
[38] M. Chen, B. Bica, L. Tüchler, A. Kann, and Y. Wang, "Statistically extrapolated nowcasting of summertime precipitation over the eastern Alps," *Adv. Atmospheric Sci.*, vol. 34, pp. 925–938, 2017.
[39] X. Zhang, T. Wang, G. Chen, X. Tan, and K. Zhu, "Convective clouds extraction from Himawari–8 satellite images based on double-stream fully convolutional networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 4, pp. 553–557, Apr. 2019.
[40] Y. Wang and B. Xiao, "Convection-UNet: A deep convolutional neural network for convection detection based on the geo high-speed imager of fengyun-4b," in *Proc. Int. Conf. Pattern Recognit., Mach. Vis. Intell. Algorithms*, 2023, pp. 163–168.
[41] J. Yang, J. Guo, H. Yue, Z. Liu, H. Hu, and K. Li, "CDNET: Cnn-based cloud detection for remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6195–6211, Aug. 2019.
[42] Q. He, X. Sun, Z. Yan, and K. Fu, "DABNET: Deformable contextual and boundary-weighted network for cloud detection in remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2021.

Renlong Hang (Member, IEEE) received the Ph.D. degree in meteorological information technology from the Nanjing University of Information Science and Technology, Nanjing, China, in 2017.

From 2018 to 2019, he was a postdoc Researcher with the Department of Computer Science and Electrical Engineering, University of Missouri-Kansas City, Kansas City, MO, USA. He is an Associate Professor with the School of Computer and Software, Nanjing University of Information Science and Technology.. He has authored or coauthored more than 40 peer-reviewed articles in international journals, such as IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. His research interests include machine learning, pattern recognition, and their applications to remote sensing image processing.

Jingquan Wang received the B.S. degree in Internet of Things engineering from the School of Computer Engineering, Bengbu University, Bengbu, China, in 2021. He is currently working toward the M.S. degree in software engineering with the School of Software, Nanjing University of Information Science and Technology, Nanjing, China.

His research interests include deep learning and convective cloud detection.

Lingling Ge received the Ph.D. degree in information and communication engineering from the Nanjing University of Information Science and Technology, Nanjing, China, in 2021.

She is currently an Engineer with the National Meteorological Information Center, Beijing, China. Her research interests include multisource data fusion, the application of machine learning on meteorology, etc.

Chunxiang Shi received the M.S. degree in atmospheric exploration from the Chinese Academy of Meteorological Sciences, Beijing, China, in 1990, and the Ph.D. degree in meteorology from the Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, in 2003.

She is currently the Chief Researcher with the National Meteorological Information Center, Beijing, China. Her research interests include multisource data fusion, the application of machine learning on meteorology, etc.

Jianfen Wei received the Ph.D. degree in climate system and climate change from the Nanjing University of Information Science and Technology, Nanjing, China, in 2018.

She is currently a Lecturer with the School of Atmospheric Science and Remote Sensing, Wuxi University, Wuxi, China. Her research interests include atmosphere-ice-ocean interactions in polar regions and AI-based sea ice prediction.