


# Detecting of Overshooting Cloud Tops via Himawari-8 Imagery Using Dual Channel Multiscale Deep Network

Shaojun Zha, Wei Jin , Caifen He, Zhiyuan Chen, Guang Si, and Zhuzhang Jin

**Abstract**—The occurrence of overshooting cloud tops can cause extremely severe weather such as strong winds and heavy rainfalls. The traditional overshooting cloud top (OT) detection methods not only need to choose a reasonable threshold, it is also very hard to make full advantage of the multispectral information of cloud images. These make small-scale OT detection very difficult with poor accuracy of OT boundary determination. In order to utilize the multispectral information of Himawari-8 satellite cloud images, in this article, we propose a method for detecting OT based on the dual channel multiscale deep network (DCMSDN). The brightness temperature of infrared window and the difference of brightness temperature between the infrared window and water vapor window are used as dual channel inputs, respectively. Then, DCMSDN introduces a multiscale prediction module to improve the accuracy of small target detection, which makes the network more suitable for the detection of the OT with small spatial scale. Experimental results indicate that the proposed method provides competitive performance with acceptable computational efficiency. Specifically, for the quantitative indicators of OTs detection, our approach achieves the accuracy of 89.36%, the precision of 95.63%, the recall of 88.90%, and the F1-measure of 91.61% for the test cloud images, which outperforms that of comparative methods.

**Index Terms**—Deep learning, dual channel multiscale deep network (DCMSDN), Himawari-8, multiscale prediction fusion, overshooting cloud tops (OT).

## I. INTRODUCTION

**O**VERSHOOTING cloud tops (OTs) are one of the atmospheric phenomena that often cause meteorological disasters. According to the Glossary of Meteorology of American Meteorological Society, an overshooting convective cloud top is defined as “a domelike protrusion above a cumulonimbus anvil, representing the intrusion of an updraft through its equilibrium

level.”[1] OT has prominent features of short life cycle, small-scale space, strong destructiveness, strong convective shear, strong updraft, etc. When a strong updraft penetrates from the top of the troposphere into the stratosphere and a deep convective storm occurs, it can cause cloud-to-ground lightning, strong winds, heavy rainfall, and other severe weather conditions, affecting the safety of flight and ground operations [2]. Therefore, it is of great significance to detect the occurrence of OT for monitoring mesoscale severe convection weather.

Utilizing the imaging characteristics of different channels of the multispectral meteorological satellite imagery, scholars have carried out the research on the OT detection. Berendes *et al* [3] proposed a method for cloud classification and detection of deep convection and OTs based on the texture of visible and near-infrared cloud images. However, the OT detection method which based on visible cloud image is not applicable during nighttime. The infrared spectral imaging channel can obtain all-weather images, and the brightness temperature of OT is often lower than that of non-OT, so the infrared window cloud images have been widely used for OT detection. The water vapor minus infrared window brightness temperature difference (WV-IRW BTD) and infrared window channel brightness temperature gradients (IRW-texture) are two typical methods. Since water vapor in the troposphere and a variety of greenhouse gases can be driven into the stratosphere by strong convection, the brightness temperature of water vapor channel in the OT is higher than that of the infrared channel. WV-IRW BTD method uses this brightness temperature difference to judge the position of OT. Martin *et al.* [4]–[8], used WV-IRW BTD method to identify the existence of OT. However, the literatures [9]–[11] pointed out that WV-IRW BTD method could easily misinterpret the convective clouds near the OT, and the effectiveness of this method depends on the water vapor conditions in the atmosphere. In order to overcome the dependence on the water vapor channel, Bedka *et al.* [1] proposed the IRW-texture method, which is based on the infrared window channel. Not only does OT show a lower brightness temperature but also its brightness temperature has a gradient difference with the surrounding due to strong convection. When the ascending force of an OT is strong, it can penetrate the tropopause into the bottom of the stratosphere and maintain a continuous cooling rate of 7–9 K/km, making them significantly cooler than the surrounding [12], [13]. Bedka *et al* [1] used the IRW-texture method combined with the tropopause temperature which obtained from the numerical weather

Manuscript received June 29, 2020; revised October 25, 2020 and November 24, 2020; accepted December 5, 2020. Date of publication December 14, 2020; date of current version January 13, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 42071323, in part by the Natural Science Foundation of Zhejiang Province under Grant LY20H180003, in part by the Natural Science Foundation of Ningbo under Grant 2019A610104, in part by the the Public Welfare Science and Technology Project of Ningbo under Grant 202002N3104, and in part by the K. C. Wong Magna Fund in Ningbo University. (Corresponding author: Wei Jin.)

Shaojun Zha, Wei Jin, Zhiyuan Chen, Guang Si, and Zhuzhang Jin are with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China (e-mail: zha950115@163.com; jinwei@nbu.edu.cn; czy273995492@hotmail.com; siguang1234@163.com; 978117901@qq.com).

Caifen He is with the Zhenghai District Meteorological Bureau, Ningbo 315202, China (e-mail: 468176571@qq.com).

Digital Object Identifier 10.1109/JSTARS.2020.3044618

prediction (NWP) to give the quantitative characteristics of OTs, and using CloudSat and Cloud Aerosol Lidar and Infrared Satellite Observation (CALIPSO) satellite data to verify the performance of OTs detection qualitatively. Wang *et al* [14] used IRW-texture combined with mesoscale analysis field data of the Local Analysis Prediction System (LAPS) to detect the position of OT, which also achieved good results. However, IRW-texture can detect the occurrence of OT, but a numerical forecast product must be combined in order to get more accurate OT location. At the same time, IRW-texture uses a fixed threshold to interpret OT, which makes it often unable to detect small-scale OT with high brightness temperature, resulting in a high-false alarm rate. Chou *et al* [15] combined WV-IRW BT method and IRW-texture method and proposed a local minimum (Local Min) method. Although the experimental results are good, the method requires visible channel in OT detection, which limits its application value at night. And similar to other mainstream methods, the method also has high requirements for threshold selection and relies on artificial knowledge and experiences.

In recent years, the methods of OT detection based on machine learning have gradually received attention. In 2017, Kim *et al* [16] distinguished the OT and non-OT regions by constructing a machine learning model, but the model classifies only individual pixels without using context information of cloud images. In 2018, based on the previous work, Kim introduced the convolutional neural network (CNN) into OT detection by using the visible channel and infrared window channel images, which was the first application of deep learning in OT detection [17]. Since the method uses visible channel image, it is only applicable to detect OTs during daytime. At the same time, during OTs detection, this method needs to chop the satellite cloud image into small patches, which leads to the problems such as difficulty in determining the OT boundary accurately and the misjudge of the anvil cloud around the OTs. In order to alleviate the difficulties faced by traditional methods, the novel OTs detection method based on dual channel multiscale deep network (DCMSDN) is proposed in this article. The proposed method can make full use of the characteristic information of OTs, which are contained in the brightness temperature of infrared window channel of Himawari-8 satellite and the brightness temperature difference between infrared window and water vapor channel. By combining the dual channel information of the cloud images with the multiscale feature prediction module, the proposed method can detect OTs accurately.

## II. DATA

### A. Imaging Channels and Physical Characteristics of the Himawari-8 Satellite

The Himawari-8 satellite is a new generation of geostationary meteorological satellite which onboard advanced Himawari imager (AHI). The AHI has 16 channels with a spatial resolution ranging from 0.5 to 2 km. The central wavelength, spatial resolution, physical properties, and applications of the AHI imaging channels are shown in Table I.

It can be seen from Table I that the remote sensing characteristics of different imaging channels of Himawari-8 satellite can

TABLE I  
SPECIFICATIONS AND PURPOSES OF HIMAWARI-8 SATELLITE CHANNELS

Channel	Himawari-8 / $\mu\text{m}$	Spatial resolution/km	Physical Properties
1	0.47	1	vegetation, aerosol
2	0.51	1	vegetation, aerosol
3	0.64	0.5	low cloud, fog
4	0.86	1	vegetation, aerosol
5	1.6	2	cloud phase
6	2.3	2	particle size
7	3.9	2	low cloud, fog, forest fire
8	6.2	2	mid-and upper level moisture
9	6.9	2	mid-level moisture
10	7.3	2	mid-and lower level moisture
11	8.6	2	cloud phase, SO <sub>2</sub> ozone content
12	9.6	2	cloud imagery,
13	10.4	2	information of cloud top cloud imagery, sea surface temperature
14	11.2	2	cloud imagery, sea surface temperature cloud top height
15	12.4	2	
16	13.3	2	

reflect the physical properties of the atmosphere from different aspects. Due to the height of OT significantly exceeding the cloud anvil, it appears as a cluster of low-brightness temperature pixels in the infrared window cloud image of 11- $\mu\text{m}$  central wavelength [18]. As for the water vapor channel, due to strong convection, the water vapor above the OT can be brought from the troposphere to the bottom of the stratosphere or even higher, so the OT often shows a higher brightness temperature in water vapor channel. Based on these differences and combined with previous experiences, Himawari-8 water vapor channel (6.2  $\mu\text{m}$ ) and infrared window channel (11.2  $\mu\text{m}$ ) data of 2-km spatial resolution were employed to detect the OTs.

### B. Dataset Construction

All experimental cloud images in this study were acquired from the Himawari-8 satellite data receiving system of our lab. The time resolution of Himawari-8 satellite data is 10 min, and the data projection method is equal latitude and longitude projection. The algorithm research area of this study is from 90 °E to 150 °E and 5 °N to 45 °N. Satellite cloud images from June to August 2019 were collected for model training and testing. During this period, severe convective weather occurred frequently in this area and OT samples were easy to obtain. In addition, the cloud images in spring at 11:00 A.M. on January 4, 2019 and 21:00 P.M. on February 23, 2019 were selected for OTs detection to verify the generalization of the model. Table II details the satellite cloud images used to construct the training and testing dataset. Referring to the date and time period in the table, we collect satellite cloud images every 30 min. An overall 1662 images were collected, of which 1500 images were randomly selected to form the training dataset, and the rest 162 images to form the test dataset. The IRW BT images and the WV-IRW BT images served as input of model. The input images were preprocessed by min-max normalization method

TABLE II  
INFORMATION OF SATELLITE IMAGERY USED FOR CONSTRUCTING DATASET

Data	Time(Beijing time)
26, June, 2019	14:30-20:00
27, June, 2019	09:30-18:30
28, June, 2019	09:30-23:30
29, June, 2019	12:00-20:30
1, July, 2019	11:00-23:30
2, July, 2019	00:00-23:00
3, July, 2019	09:30-23:30
4, July, 2019	00:00-21:30
5, July, 2019	14:30-23:30
6, July, 2019	06:00-12:30
9, August, 2019	18:30-23:30
10, August, 2019	00:00-05:30
11, August, 2019	10:00-18:00

[17], which can be described as  $(X-X_{\min})/(X_{\max}-X_{\min})$ , where  $X_{\min}$  and  $X_{\max}$  are the minimum and maximum pixel value of the input image, while  $X$  represents the current pixel value.

### III. METHODOLOGY

#### A. Analysis of the Characteristics of OT

OT is also called penetrating convective cloud. When OT appears, it is often accompanied by deep convection, which makes it easy to be confused with deep convective cloud during OT detection. Two common OT detection methods are depicted below.

The WV-IRW BTD method is proposed based on the BT value, which uses the brightness temperature differences for OT detection. The literature [1] pointed out the premise of this technique for OT detection: 1) The atmospheric temperature profile warms with height in the lower stratosphere, 2) water vapor is forced into the lower stratosphere at levels above the physical cloud top by the overshooting storm updraft, 3) this water vapor emits at the warmer stratospheric temperature whereas emission in the infrared window channel originates from the colder physical cloud top, and 4) positive differences between the warmer water vapor brightness temperature (WV BT) and colder IRW BT can therefore identify where overshooting is occurring. However, the maximum brightness temperature difference area does not always correspond to OT. This is because the source of water vapor in the stratosphere may not be related to OT.

The IRW-texture method is developed based on the difference between the brightness temperature of OT and the average brightness temperature of the surrounding anvil cloud. Since it does not depend on the water vapor distribution in the lower stratosphere, IRW-texture method addresses the dependence of the WV-IRW BTD method on water vapor, thus overcomes the over-detection of OT by the WV-IRW BTD method. The IRW-texture method includes the following steps: 1) Define candidate OT pixels (IRW BT below the bright temperature threshold  $T_0$ ); define anvil cloud pixels (noncandidate OT pixels and IRW BT below the bright temperature threshold  $T_1$ , where  $T_0 < T_1$ ); 2) taking a single candidate OT pixel as the center, the average brightness temperature  $\bar{T}_1$  of the anvil cloud pixels in a certain

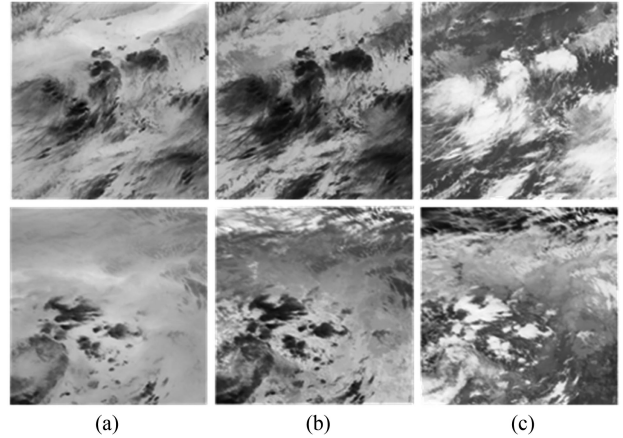


Fig. 1. Satellite cloud images of different imaging channels. (a) WV BT images. (b) IRW BT images. (c) BTD images.

range is calculated; 3) compare the brightness temperature of a single candidate OT pixel with the corresponding average brightness temperature of the anvil-shaped cloud pixels. If the brightness temperature of candidate OT pixel is  $T$  K lower than the average brightness temperature of anvil-shaped cloud (where  $T$  is a threshold, determined by experience), the candidate OT pixel can be confirmed as real OT pixel, repeat this step 2) and 3) for each candidate OT pixel to get OT region. It can be seen from the above steps that the method is greatly affected by the thresholds and is easily affected by some subjective factors. According to the different characteristics of WV-IRW BT and IRW-texture methods, combining with CNN, a novel OTs detection method is proposed in this article. The WV BT images, IRW BT images, and BTD images at specific times are shown in Fig. 1.

In Fig. 1, on the brightness temperature images, the darker the WV BT and IRW BT images are the lower the brightness temperature will be, and vice versa. At the same time, textures of IRW BT images are more obvious than that of WV BT images, which can be beneficial for CNN learning. BTD images show the brightness temperature difference, the brighter the larger the difference. Since most OTs regions come in small size and usually blurred with the deep convective clouds, their boundaries cannot be defined accurately. This leads to a large number of false detections in OTs detection by traditional methods. Therefore, rich image features and accurate positioning are needed in OTs detection, and clear boundaries also must be provided at the same time.

In recent years, U-net has been successfully applied in image segmentation, especially in medical image segmentation [19]. By end-to-end training from very few images, U-net can obtain accurate target boundary location in image segmentation. The U-net consists of two paths: Down-sampling and up-sampling. The down-sampling is used to obtain image information. The up-sampling fully acquires multilevel features of the image and simultaneously achieves target positioning. Since OTs detection can be regarded as an image segmentation problem, U-net can help to achieve OTs detection. However, the pooling operation of down-sampling in the U-net will lead to the loss of some

information of cloud image, which not only makes it difficult to accurately determine the OT area but also causes some small-scale OT missed detection due to the constant change of cloud system.

To reduce the disadvantages of pooling operation, the feature pyramid model [20] which has been widely used in target detection can be introduced to the U-net structure. Generally speaking, in deep network, the low-level feature maps can depict target location information more accurately while the high-level feature maps only contain rough location information of the target; in terms of reflecting the semantic features of the target, the high-level feature maps are more meaningful. By predicting and fusing the multiscale feature maps to obtain higher resolution features, the feature pyramid mode can detect targets with different scales and improve the detection accuracy of small targets. Moreover, U-net can just process single channel cloud image which cannot reflect the occurrence and development of OTs comprehensively, thus restricting its performance in OTs detection.

Following the analysis mentioned above, a DCMSDN is proposed for detecting OTs by the IRW BT image and BTD image. Due to introducing dual channel and multiscale structure, inheriting the down-sampling and up-sampling paths of U-net, the DCMSDN can take full advantages of the multispectral characteristics of satellite cloud images and has good OT boundary positioning capability.

### B. Structure of the Dual Channel Multiscale Deep Network

The single-channel input structure of the traditional U-net network is difficult to comprehensively utilize the multispectral information of the satellite cloud images. Based on the codec structure of the U-net, starting from the actual needs of OTs detection, the U-net has been improved as follows: 1) Designing a coding structure with dual-channel input. The 2-D convolution and down-sampling of the IRW BT images and BTD images, which are suitable for the detection of OTs, are used as the inputs of the two coding channels. After the encoding procedure, the information of the two channels will be fused by superimposing the corresponding scale feature maps of the two channels. In the decoding procedure, corresponding to the encoding operation, the up-sampling and convolution operation were carried out on the fused feature map to gradually recover the size of each hidden layer, so as to restore the details of OT and output the probability maps to position the OT accurately; 2) introducing a multiscale prediction module. Due to the small spatial scale of OT, the introduction of multiscale prediction module in the coding process can fuse the features of cloud maps at different scales, which not only enhances the detection accuracy for small-scale OT but also guides the network to carry out OT detection oriented learning; 3) using batch normalization (BN) [21] in training. Except for the last convolutional layer, there is an additional BN layer after each convolutional layer. This operation can effectively reduce the problem of gradient vanishing of low-level neural network during error back propagation. At the same time, in order to reduce the loss of OT boundary information during down-sampling and ensure the consistent size of input and output

images, the padding parameters of convolutional and pooling were set as 1, and then the Crop layer was replaced with a Copy layer in the feature fusion; 4) introducing inception module. Inspired by Inception V3 [22], in order to control the number of parameters while increasing the depth and width of the network, which are aimed to extract the spectral information of the cloud images more effectively,  $1 \times 3$  convolutional layer and then followed by  $3 \times 1$  convolutional layer instead of the traditional  $3 \times 3$  convolutional layer are implemented. This not only reduces the number of network parameters but also improves the training efficiency. The structure of DCMSDN is shown in Fig. 2.

It can be seen, the DCMSDN has a symmetric structure. In the stage of down-sampling, DCMSDN applies same levels of Group module A and Max pooling module processing to the dual-channel inputs IRW BT image and BTD image, respectively, so that OTs related characteristics can be obtained. Every step in the up-sampling path consists of a Predict module and then concatenation of the feature maps of the dual-channel contracting paths followed by a Group module B. After the same levels of processing as the down-sampling, the multispectral, multilevel, and multiscale characteristics of OTs can be fused. At the final layer a  $1 \times 1$  convolution is used to map the feature maps, which have been processed by the fourth layer group module B, to the results and obtain the desired OTs. In order to training DCMSDN effectively, a reasonable loss function is designed by calculating the distance between the output image and the real label image.

Group module A and group module B are shown in Fig. 3. Both of them consist of the repeated application of two  $1 \times 3$  and  $3 \times 1$  convolutions, each  $3 \times 1$  convolutions is followed by a batch normalization (BN) and rectified linear unit (ReLU). The design inspiration of these structures comes from Inception V3, and the network parameters can be effectively reduced by convolution kernel decomposition. For example, a  $3 \times 3$  convolution kernel with 9 parameters is decomposed to one  $1 \times 3$  and one  $3 \times 1$  convolution kernel, and the parameters can be reduced to 6. At the same time, the convolution kernel is decomposed in this way, which greatly enhances the nonlinearity of the network.

The prediction module is shown in Fig. 4. The module designed draws on the feature pyramid model. For input feature maps of each scale, a  $1 \times 1$  convolution is used to map feature maps to the desired prediction result of the corresponding scale. Then, the input feature maps and prediction result is fused by concatenating correspondingly using the skip connection. In order to eliminate the aliasing effect,  $1 \times 3$  and  $3 \times 1$  convolutions are applied to the fused results so as to obtain output features of the corresponding layer.

As for the design of loss function, considering that in the OTs detection, it is not uncommon that the OTs occupies only a very small region of the satellite cloud image. This often causes the learning process to get trapped in local minima of the loss function yielding a network whose predictions are strongly biased toward background. As a result the OTs regions are often missing or only partially detected. The traditional way to solve this problem is often to resort to loss functions based on sample reweighting, where foreground regions are given more

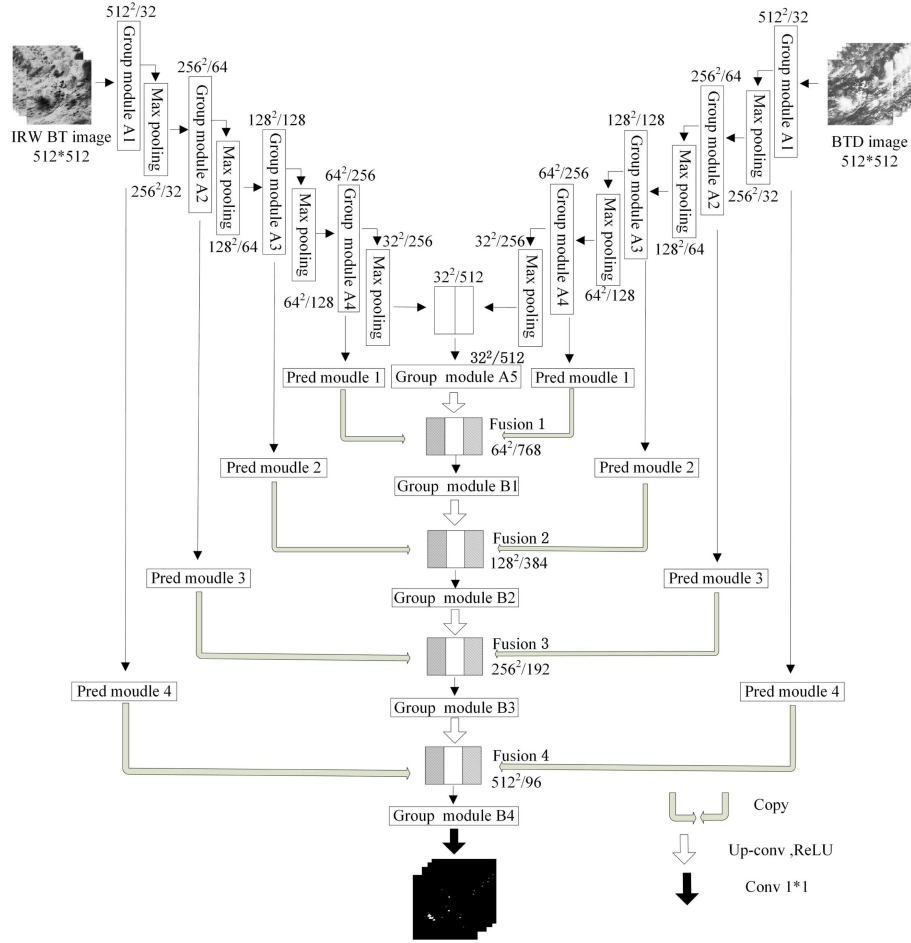


Fig. 2. Structure of DCMSDN.

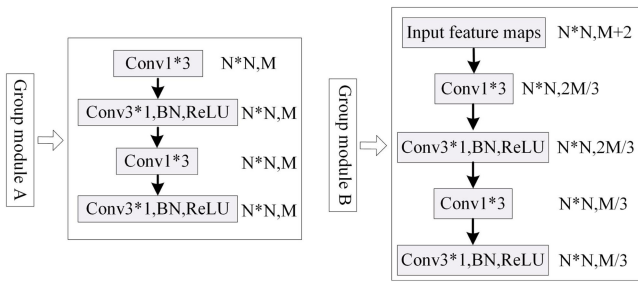


Fig. 3. Structure of the Group module.

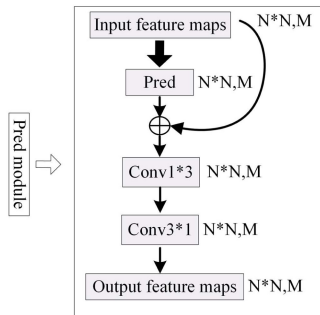


Fig. 4. Structure of the Predict module.

importance than background ones during learning. In this work we propose a novel loss function based on the Dice similarity coefficient which can be defined as

$$S = \frac{2|X \cap Y|}{|X| + |Y|} \quad (1)$$

where the  $X$  represents the predicted OTs regions and the  $Y$  represents the real OTs regions. Using this formulation we do not need to assign weights to samples of different classes to establish the right balance between OTs and non-OTs pixels, and we can obtain much better results than that of the traditional methods.

### C. Determination of the Sample Label of OTs

The accuracy and reliability of the sample labels are critical to the OTs detection. Literature [17] proposes to divide the satellite cloud image into patches with the size of  $31 \times 31$ , and mark the image patches as OT and non-OT (1 indicates OT, and 0 indicates a non-OT). Due to the irregular boundary of OT, by marking the entire patches as 1 or 0, this method is not only difficult to define the boundary of OT accurately but also often marks the non-OT pixels surrounding the OT region as OT by mistake. Based on Chou's method [15], we propose a pixel-level

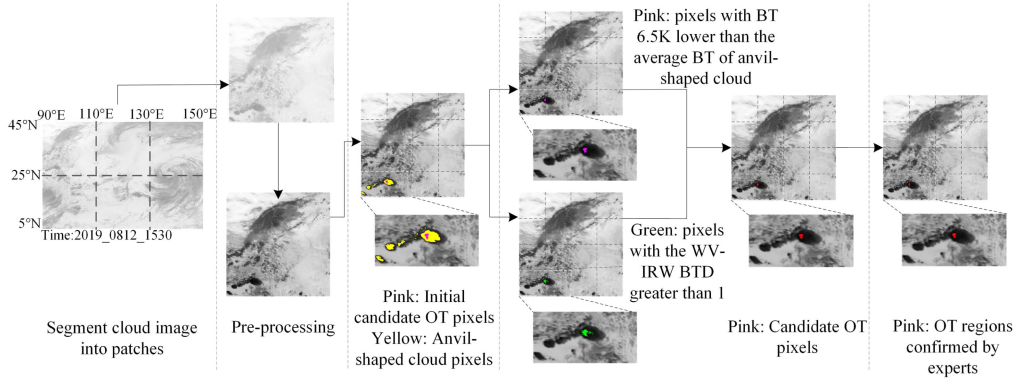


Fig. 5. Process flow diagram of labeling OTs regions.

classification method to obtain the OTs labels as the reference data. The specific steps are as follows.

- 1) Segmenting the infrared window brightness temperature (IRW BT) cloud image into patches.
- 2) Finding the minimum brightness temperature  $T_{\min}^i$  of the  $i$ th patch.
- 3) Determining the initial candidate OT pixels  $T_{H\_OT(n)}^i$  which satisfy both  $T_{H\_OT(n)}^i < T_{\min}^i + 4$  and  $T_{H\_OT(n)}^i < 215$ , where  $n = 1, 2, \dots, N$ .
- 4) Determining the anvil-shaped cloud pixels  $T_{Z\_OT(m)}^i$  which satisfy both  $T_{Z\_OT(m)}^i < T_{\min}^i + 15$  and  $T_{Z\_OT(m)}^i < 225$ , where  $m = 1, 2, \dots, M$ .
- 5) With each initial candidate OT pixel as the center, calculate the average brightness temperature of all  $T_{Z\_OT(m)}^i$  within 8–24 km and denoted as  $\bar{T}^i$ .
- 6) The  $T_{H\_OT(n)}^i$  can be determined as candidate OT pixel, if  $T_{H\_OT(n)}^i < \bar{T}^i - 6.5$  and the WV-IRW BTD of the corresponding pixel is greater than 1.
- 7) Repeat steps 3), 4), 5), and 6) to obtain all candidate OT pixels, and then merge adjacent OT pixels to form a candidate OTs regions.
- 8) The candidate OTs regions were further confirmed by two weather experts with more than 5-year experience in cloud image interpretation. Through the above steps, almost all OTs have been detected, but some anvil clouds with lower brightness temperature were mistakenly classified as OTs, and there is a certain false alarm. When confirming the candidate OTs regions, weather experts will refer to local meteorological data such as actual precipitation to eliminate the false alarm pixels, so as to objectively delineate the boundary of each OT.

(OT tends to correspond to the region where heavy precipitation occurs). And if the opinions of the two experts are inconsistent, the delineated results were evaluated by another meteorologist with more than 10-year experience in cloud image interpretation to make sure they agree. Based on the above scheme, we mark the OTs regions and determine each OT boundary. The process flow diagram of labeling OTs regions is shown in Fig. 5.

Fig. 6 shows the IRW BT images and the BTD images at 10:30 on June 27 and 15:30 on August 12, 2019 as well as the marked OTs regions.

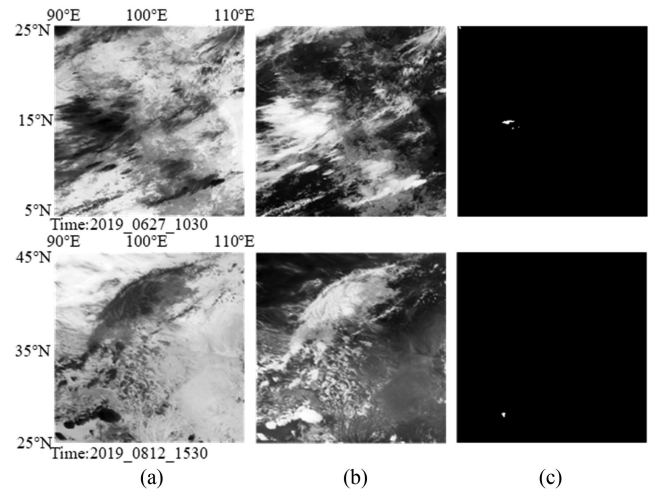


Fig. 6. IRW BT images, the BTD images and the marked OTs regions. (a) IRW BT images. (b) BTD images. (c) Marked OTs regions.

#### D. Performance Indicators of OTs Detection

The performance indicators of evaluating different OTs detection methods were generated based on true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Where TP is the number of pixels that belong to OT and were detected exactly as OT by the model; TN is the number of pixels that do not belong to OT and were assigned to non-OT correctly; FP is the number of pixels that do not belong to OT but were predicted positively to OT while FN are the pixels that belong to OT but were not detected. Specifically, the five performance indicators were calculated as follows.

- 1) *Accuracy*: This indicates the ratio of correctly assigned pixels in the total pixels as seen in (2). This indicator can be used to measure the overall detection accuracy of the model

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (2)$$

- 2) *Precision*: This indicates the ratio of correctly detected as OT pixels in the total detected OT pixels as seen in (3). The higher value of precision, the less likely the non-OT

pixels will be detected as OT pixels

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (3)$$

- 3) *Recall*: The recall is the ratio of the correctly detected as OT pixels in the total true OT pixels as indicated in (4). The higher value of recall, the less the OT pixels will be missed detection

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (4)$$

- 4) *F1-Measure*: Since recall and precision are contradictory in some cases, high recall often means low precision, so F1-measure is included in the evaluation indicators. F1-measure is a weighted harmonic mean of precision and recall, computed using (5). The higher value of F1-measure, the better the comprehensive performance of recall and precision, and the better the ability of OTs detection

$$\text{F1-Measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (5)$$

- 5) *False Alarm Ratio (FAR)*: The FAR, gives the fraction of detected as OT pixels that were observed to be non-OT pixels as seen in (6). Contrary to precision, the lower the FAR value, the less likely the non-OT pixels will be mistakenly detected as OT pixels

$$\text{FAR} = \frac{\text{FP}}{\text{TP} + \text{FP}}. \quad (6)$$

#### IV. EXPERIMENTAL RESULTS

The performance of the proposed OTs detection method was evaluated by the following experiments. All the experiments were performed on an Intel Core i7-7700 CPU at 3.5 GHz with 64-GB RAM. Unless otherwise specified, for all models, we initialized the weights with random values, set the batch size to 5, set learning rate to 0.0001 and trained 2000 epochs on 11G NVIDIA GeForce GTX 1080Ti GPU, stochastic gradient descent (SGD) as optimizer.

##### A. Visualization of Feature Information

In order to evaluate the effectiveness of the DCMSDN proposed in this article, DCMSDN feature maps at different levels are visualized and compared against the single-channel feature maps of the U-net. Fig. 7 shows the visualization feature maps of the group module A1–A4 from DCMSDN, cloud images occurred at 19:30 on June 26, 2019 as an example.

It can be seen that in the contracting path (the path of down-sampling) of the DCMSDN, the feature maps of the IRW BT image mainly depict the areas with low brightness temperature (dark areas in the input image), the characteristics of the areas with relatively low bright temperature are extracted by the shallow group module and with the deepening of the network, OT can be identified clearly. For the feature maps of BT image, the contracting path extract the overall contour and texture features of BT image, with the deepening of the network, the areas with large bright temperature differences can be well represented.

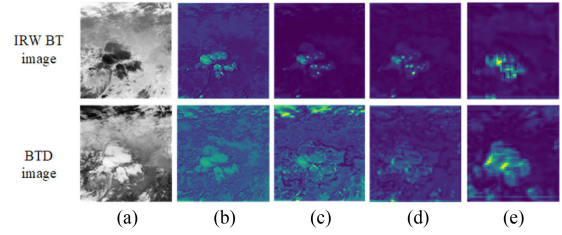


Fig. 7. Feature maps of group module A1–A4. (a) Input images. (b) Feature maps of group module A1. (c) Feature maps of group module A2. (d) Feature maps of group module A3. (e) Feature maps of group module A4.

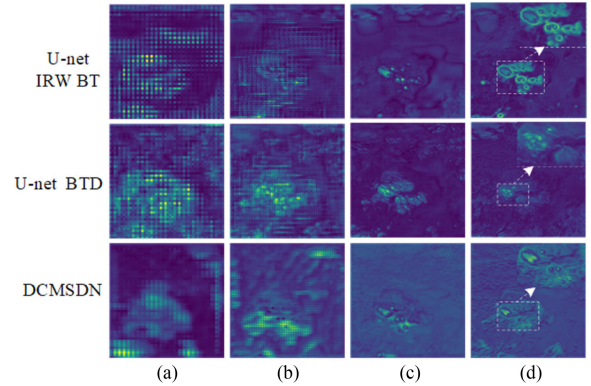


Fig. 8. Comparison of the fusion feature maps of different levels between the DCMSDN and the single channel U-net. (a) Fusion 1. (b) Fusion 2. (c) Fusion 3. (d) Fusion 4.

Every step in the expansive path (the path of up-sampling) fuses the corresponding multiscale feature map which from the contracting paths of IRW BT image and BTD image by concatenating. Due to the features of IRW BT image and BTD image were extracted by the down-sampling process of DCMSDN, this structure is conducive to guiding the network to carry out fine characterization of the OT region, thus improving the accuracy of OTs detection. Fig. 8 shows the comparison of the fusion feature maps of different levels between the DCMSDN and the single channel U-net.

The different levels feature maps of the IRW BT image and BTD image which represented by the U-net are shown in the first and the second row of the Fig. 8, while the corresponding feature maps which are obtained by the DCMSDN via fusion the information of the IRW BT image and BTD image are shown in the third row of the Fig. 8. It can be seen that the single-channel network like U-net cannot represent the characteristic of OT well; consequently it is easy to mistakenly identify the deep convective clouds surrounding the OT region as OT. On the other hand, by fusing the dual-channel and multiscale features of IRW BT image and BTD image, the DCMSDN can not only detect small-scale OT effectively but also better describe the boundary of OT well and distinguish OT from deep convective clouds simultaneously.

##### B. Effectiveness Analysis of the Dual Channel Model

In order to evaluate the effectiveness of DCMSDN for OTs detection, we took BTD images and IRW BT images as U-net

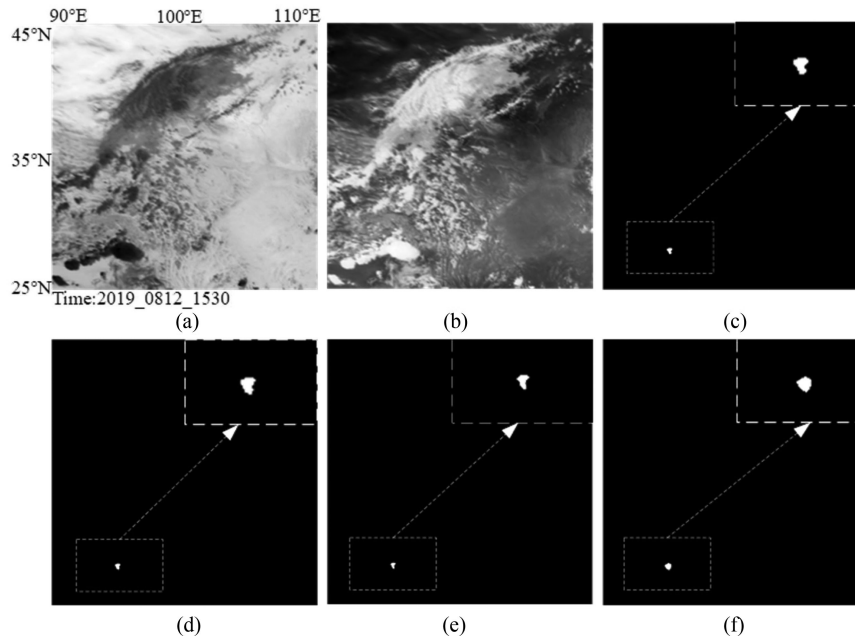


Fig. 9. Comparison of OT detection results via single-channel and dual-channel networks, the OTs are denoted as white in the detection results. (a) IRW BT image. (b) BT image. (c) Labeled OT image. (d) Detection result of the DCMSDN. (e) Detection result of the U-net\_IRW. (f) Detection result of the U-net\_BTD.

TABLE III  
COMPARISON OF EXPERIMENTAL RESULTS

Model	Accuracy /%	Precision /%	Recall /%	F1-measure /%	FAR /%
U-net_BTD	77.84	86.84	73.81	78.84	13.16
U-net_IRW	84.46	88.43	80.76	84.42	11.57
DCMSDN	89.36	95.63	88.90	91.61	4.37

inputs for OTs detection, and then compared the detection results with the results of DCMSDN. Table III tabulates the comparison of experimental results. Among them, U-net\_BTD and U-net\_IRW represent the single-channel detection results using the BT image and IRW BT image as the U-net input. In contrast, the DCMSDN represents the detection results using both the BT images and IRW BT images as the dual channel inputs. Compared with the results obtained by single-channel U-net, the outputs of DCMSDN are about 5% higher at least in term of Accuracy and F1-measure, and more than 7% in term of Precision and Recall.

In order to show the detection effect visually, Fig. 9 shows the detection results of various models (the upper right corner is the zoom in of the detection result). From visual assessment, the single-channel models U-net\_IRW and U-net\_BTD can detect the rough region of the OT but produce inaccurate definition of the OT boundary. In contrast, the detection result of DCMSDN has obvious advantages in depicting OT boundary by combining feature maps of different levels of IRW BT image and BT image.

### C. Comparisons With Other Methods

Various techniques and methods have been used for detecting OTs in satellite imagery. Each technique has its own respective

TABLE IV  
COMPARISON OF OTS DETECTION ACCURACY VIA DIFFERENT METHODS

Model	Accuracy /%	Precision /%	Recall /%	F1-measure /%	FAR /%
IRW-texture	63.30	100	40.61	56.24	0
WV-IRW BTD	70.79	82.13	69.01	71.80	17.87
Local Min	75.23	96.04	66.87	84.42	3.96
DCMSDN	89.36	95.63	88.90	91.61	4.37

accuracy level. Here, DCMSDN was compared with traditional methods IRW-texture [14], WV-IRW BTD [15], and Local Min [15].

Table IV tabulates the comprehensive evaluation indicators of OTs detection by different methods. It can be seen that the Precision of the IRW-texture method reaches 100%, but its Recall is very low. The main reason is that the threshold of the bright temperature in the IRW-texture is a sensitive factor, which has a particularly important influence on the detection results. For example, some OTs may easily be missed with a low threshold, while a higher threshold is prone to label some convective clouds, especially deep convective clouds, as OTs and resulting in many misjudgments. Although the WV-IRW BTD method is not sensitive to the threshold, it is affected greatly by water vapor. Meanwhile, the selection of the threshold for the brightness temperature difference is also very important, so there are a lot of false detections. Compared with the above two methods, the Local Min method combines the advantages of the above two methods and improves the accuracy of OTs detection. However, the detection results of the Local Min are not as good as the method proposed in this article. From the perspective of Accuracy and F1-measure, the proposed method is about 13% and 7% higher than that of the Local Min method, respectively.



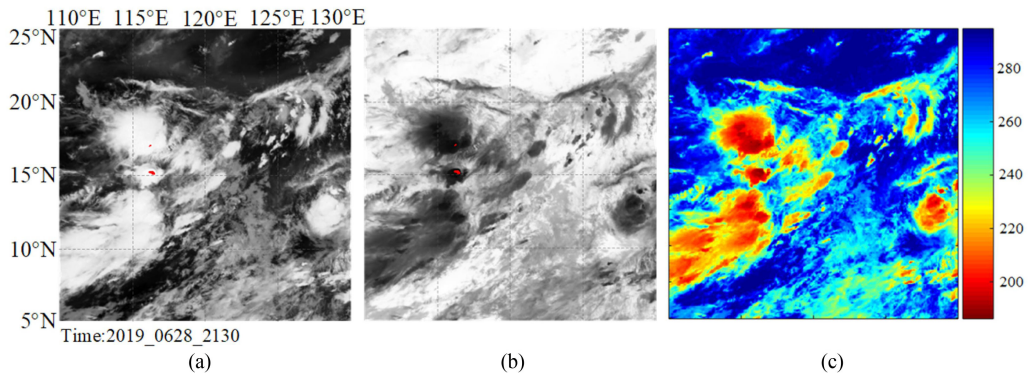


Fig. 10. Visualization of the OTs, the OTs are denoted as red in (a) and (b). (a) OTs on the BT image. (b) OTs on the IRW BT image. (c) Pseudocolor visualization of the IRW BT image.

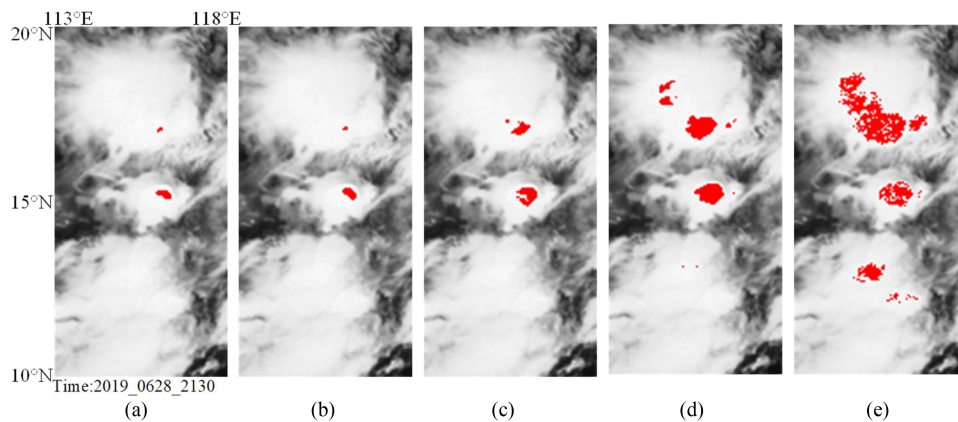


Fig. 11. Comparison of OTs by different methods, the OTs are denoted as red. (a) Label image. (b) DCMSDN. (c) Local Min. (d) IRW-Texture. (e) WV-IRW BT image.

From the perspective of FAR, the proposed method DCMSDN is not the best one, but considering all these evaluation indicators, the MCMSDN is the optimal method.

Fig. 10 shows the cloud images at 21:30 on June 28 as an example, Fig. 10(a) and (b) are the BT image and IRW BT image, respectively, the OTs regions are shown in red. It can be seen that the OT located at the low brightness temperature region of IRW BT image and the high water vapor-infrared window brightness temperature difference region of BT image. However, not all pixels in the low brightness temperature region belong to OT. Fig. 10(c) is the pseudocolor visualization of the Fig. 10(b), from which it can be seen clearly the region of OT is much smaller than the low brightness temperature region. The zoom in of the OTs regions of Fig. 10 were shown in Fig. 11, and the results of OTs detection by different methods were also shown in Fig. 11, where Fig. 11(a) was the labeled image which was confirmed by weather experts, and Fig. 11(b)–(e) were the results of OTs detection using DCMSDN, Local Min, IRW-Texture, and WV-IRW BT image-based methods, respectively. It can be seen from the Fig. 11 that lots of pixels that do not belong to OTs were assigned as OTs mistakenly for the IRW-texture and WV-IRW BT image-based methods. The performance of the Local Min-based method is better than that of the IRW-texture-based method and the WV-IRW BT image-based method, but the false alarm of OTs still

exists. According to the discussions above, the proposed method which based on DCMSDN is the best in terms of detection accuracy and visual inspection.

#### D. Analysis of the Generalization Ability

In order to verify the generalization ability of the DCMSDN, we have selected two cloud images in spring which at 11:00 A.M. on January 4, 2019 and 21:00 P.M. on February 23, 2019 for OTs detection. The test results as shown in Fig. 12. It can be seen that the OTs areas detected by the DCMSDN are coincided with the label generally. Table V gives quantitative evaluation indicators. As seen from the table, most of the indicators that measure the accuracy of OTs detection exceeded 80% and the FAR less than 20%. Judging by these indicators, the model constructed in this article is suitable for OTs detection in other seasons, and has strong generalization performance.

#### E. Analysis of Computational Efficiency

To evaluate the computation efficiency, for 162 cloud images in the test dataset, the total time of OTs detection by various methods was counted respectively. And then, we calculate the average time of OTs detection for single cloud image of each method. Table VI tabulates the average computing time for

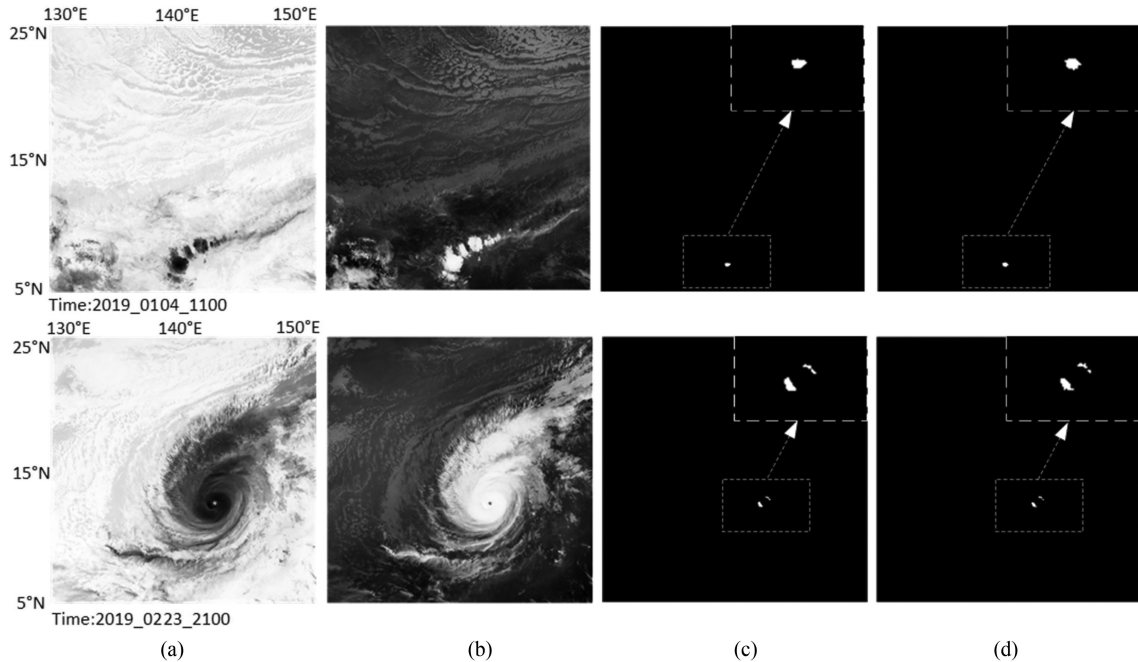


Fig. 12. OTs detection results of another season, the OTs are denoted as white in the detection results. (a) IRW BT images. (b) BT-D images. (c) The detection results of the DCMSDN. (d) The labeled OTs images.

TABLE V  
QUANTITATIVE EVALUATION INDICATORS FOR OTS  
DETECTION IN OTHER SEASONS

Cloud images at different times	Accuracy /%	Precision /%	Recall /%	F1-measure /%	FAR /%
11:00 a.m. on January 4, 2019	81.25	92.11	79.55	85.37	7.89
21:00 p.m. on February 23, 2019	83.95	81.63	90.91	86.02	18.37

TABLE VI  
AVERAGE COMPUTATION TIME TO DETECT OTS WITH DIFFERENT MODELS

Model	Time/s
IRW-texture	0.374
WV-IRW BT-D	0.579
Local Min	0.889
U-net	0.045
DCMSDN	0.130

IRW-texture, WV-IRW BT-D, Local Min, U-net, and DCMSDN methods. It can be seen that Local Min takes the longest time to detect a single cloud image as of 0.889 s, followed by WV-IRW BT-D and IRW-texture, 0.579 and 0.374 s, respectively. The average OTs detection time of the U-net-based method is the shortest, requiring only 0.045 s. Because the structure of DCMSDN is more complex than that of U-net, the average computing time of the proposed DCMSDN method is slightly longer than that of the U-net method. In general, compared with

the competing methods, the proposed method has higher OTs detection performance with acceptable efficiency.

## V. CONCLUSION

This article studies the deep learning-based solutions to OTs detection for the Himawari-8 satellite imagery and constructs a new dual channel multiscale deep network as DCMSDN in detail. The DCMSDN follows the codec structure of the U-net and has dual-channel input. The features of OTs region can be extracted from IRW BT image and BT-D image, respectively, by using the dual-channel structure of DCMSDN. By introducing multiscale prediction module and inception module, and adopting the Dice similarity coefficient-based loss function, the DCMSDN is more suitable for the detection of OT with small spatial scale. Different OTs detection methods were benchmarked on their detection accuracy by data labeled by weather experts, the results show that the proposed method which is based on DCMSDN is better than U-net-based methods and other traditional methods such as IRW-texture, WV-IRW BT-D, and Local Min etc. Future works will aim at enhancing the generalization ability of the model to further improve the accuracy of OTs detection in different seasons, and extending the DCMSDN to other satellite cloud image processing tasks.

## REFERENCES

- [1] K. Bedka *et al.*, "Objective satellite-based detection of overshooting tops using infrared window channel brightness temperature gradients," *J. Appl. Meteorol. Climatol.*, vol. 49, no. 2, pp. 181–202, Feb. 2010.
- [2] D. W. Reynolds, "Observations of damaging hailstorms from geosynchronous satellite digital data," *Monthly Weather Rev.*, vol. 108, no. 3, pp. 337–348, Mar. 1980.

- [3] T. A. Berendes *et al.*, "Convective cloud identification and classification in daytime satellite imagery using standard deviation limited adaptive clustering," *J. Geophys. Res.*, vol. 113, Oct. 2008, Art. no. D20207.
- [4] D. W. Martin *et al.*, "Over-ocean validation of the global convective diagnostic," *J. Appl. Meteorol. Climatol.*, vol. 47, no. 2, pp. 525–543, Feb. 2008.
- [5] S. Fritz and I. Laszlo, "Detection of water vapor in the stratosphere over very high clouds in the tropics," *J. Geophys. Res.*, vol. 98, no. 12, pp. 22959–22967, Dec. 1993.
- [6] S. A. Ackerman, "Global satellite observations of negative brightness temperature differences between 11 and 6.7  $\mu\text{m}$ ," *J. Atmos. Sci.*, vol. 53, no. 19, pp. 2803–2812, Oct. 1996.
- [7] J. Schmetz *et al.*, "Monitoring deep convection and convective overshooting with METEOSAT," *Adv. Space Res.*, vol. 19, no. 3, pp. 433–441, Mar. 1997.
- [8] 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," *Atmos. Res.*, vol. 83, no. 2, pp. 505–518, 2007.
- [9] K. M. Bedka *et al.*, "Validation of satellite-based objective overshooting cloud-top detection methods using CloudSat cloud profiling radar observations," *J. Appl. Meteorol. Climatol.*, vol. 51, no. 10, pp. 1811–1822, Oct. 2012.
- [10] M. Setvák *et al.*, "A-Train observations of deep convective storm tops," *Atmos. Res.*, vol. 123, pp. 229–248, Apr. 2013.
- [11] K. M. Bedka and K. Khlopenkov, "A probabilistic multispectral pattern recognition method for detection of overshooting cloud tops using passive satellite imager observations," *J. Appl. Meteorol. Climatol.*, vol. 55, no. 9, pp. 1983–2005, May 2016.
- [12] A. J. Negri, "Cloud-top structure of tornado storms on 10 April 1979 from rapid scan and stereo satellite observations," *Bull. Amer. Meteorol. Soc.*, vol. 63, no. 10, pp. 1151–1859, Nov. 1982.
- [13] R. F. Adler *et al.*, "Thunderstorm top structure observed by aircraft overflights with an infrared radiometer," *J. Climate Appl. Meteorol.*, vol. 22, no. 4, pp. 579–593, Apr. 1983.
- [14] X. Wang, Y. Huang, and L. Guan, "Satellite-based detection of overshooting cloud tops using infrared window channel brightness temperature," *J. Meteorol. Sci.*, vol. 33, no. 1, pp. 71–76, Feb. 2013.
- [15] C. B. Chou, "Methods and experiments of detecting overshooting top by using Himawari-8 imagery," *Adv. Geosci.*, vol. 6, no. 3, pp. 190–200, Jun. 2016.
- [16] M. Kim *et al.*, "Detection of tropical overshooting cloud tops using Himawari-8 imagery," *Remote Sens.*, vol. 9, Jul. 2017, Art. no. 685.
- [17] M. Kim, J. Lee, and J. Im, "Deep learning-based monitoring of overshooting cloud tops from geostationary satellite data," *GISci. Remote Sens.*, vol. 55, no. 5, pp. 763–792, May 2018.
- [18] Y. Liu, J. Xia and C. X. Shi, and Y. Hong, "An improved cloud classification algorithm for China's FY-2C multi-channel images using artificial neural network," *Sensors*, vol. 9, no. 7, pp. 5558–5579, Jul. 2009.
- [19] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2015, pp. 234–241.
- [20] T. Lin *et al.*, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 936–944.
- [21] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate Shift," in *Proc. 32nd Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
- [22] C. Szegedy *et al.*, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2818–2826.



**Shaojun Zha** received the bachelor's degree in communication engineering from Hehai University Wenhua College, Maanshan, China, in 2017. She is currently working toward the master's degree in electronics and communication engineering with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, China.

Her research interests include image segmentation, remote sensing image understanding and application, deep learning, and pattern recognition.



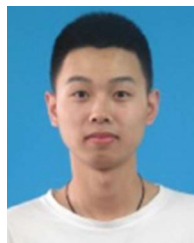
**Wei Jin** received the Ph.D. degree in optical engineering from Chongqing University, Chongqing, China, in 2006.

He is currently a Professor with the faculty of electrical engineering and computer science, Ningbo University, Ningbo, China. His research interests include sparse representation, deep learning, computer vision, and image processing.



**Caifen He** received the master's degree in atmospheric sciences from Lanzhou University, Lanzhou, China, in 2011.

She is currently a Weather Forecaster with Zhenhai District Meteorological Bureau, Ningbo, China. Her research interests include satellite cloud image processing, weather radar data analysis, and deep learning.



**Zhiyuan Chen** received the bachelor's degree in communication engineering from Binzhou University, Binzhou, China, in 2018. He is currently working toward the master's degree in electronics and communication engineering with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, China.

His research interests include remote sensing, image segmentation, and deep learning.



**Guang Si** received the bachelor's degree in computer engineering from Ludong University, Ludong, China, in 2018. He is currently working toward the master's degree in electronics and communication engineering with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, China.

His research interests include meteorological signal processing and deep learning.



**Zhuzhang Jin** received the bachelor's degree in computer engineering in Ningbo University of Technology, Ningbo, China, in 2019. He is currently working toward the master's degree in computer science and technology with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, China.

His research interests include deep learning, saliency detection, and its application in remote sensing target detection.