A Priori Land Surface Reflectance Synergized With Multiscale Features Convolution Neural Network for MODIS Imagery Cloud Detection

Nan Ma [ORCID], Lin Sun [ORCID], Chenghu Zhou, Yawen He, Chuanxiang Dong, Yu Qu, and Huiyong Yu

Abstract—Moderate resolution imaging spectrometer (MODIS) images are widely used in land, ocean, and atmospheric monitoring, due to their wide spectral coverage, high temporal resolution, and convenient data acquisition. Accurate cloud detection is critical to the fine processing and application of MODIS images. Owing to spatial resolution limitations and the influence of mixed pixels, most MODIS cloud detection algorithms struggle to effectively recognize of clouds and ground objects. Here, we propose a novel cloud detection method based on land surface reflectance and a multiscale feature convolutional neural network to achieve high-precision cloud detection, particularly for thin clouds and clouds over bright surface. A monthly surface reflectance dataset was constructed by MODIS products (MOD09A1) and employed to provide background information for cloud detection. Difference-based samples were obtained using surface reflectance as well MODIS images of different phases based on difference operations. The multiscale feature network (MFCD-Net) using an atrous spatial pyramid pooling and a channel and spatial attention module integrated low-level spatial features and high-level semantic information to capture multiscale features and generate a high-precision cloud mask. For cloud detection experiments and quantitative analysis, 61 MODIS images acquired at different times on various underlying surface types were used. Cloud detection results were compared to those of UNet, Deeplabv3+, UNet++, PSPNet, and top of atmosphere-based (MFCD-TOA) methods. The proposed method performed well, with the highest overall accuracy (96.55%), precision (92.13%), and recall (88.90%). It improved cloud detection accuracy in various scenarios, reducing thin cloud omission and bright surface misidentification.

Index Terms—Cloud detection, difference-based samples, land surface reflectance (LSR) dataset, moderate resolution imaging spectrometer (MODIS), multiscale feature.

Nan Ma is with the School of Geoscience, China University of Petroleum, Qingdao 266580, China (e-mail: 552068321@qq.com).

Lin Sun, Chuanxiang Dong, Yu Qu, and Huiyong Yu are with the College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China (e-mail: sunlin6@126.com; 281710395@qq.com; 1916954123@qq.com; 603983886@qq.com).

Chenghu Zhou is with the State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China (e-mail: zhouch@lreis.ac.cn).

Yawen He is with the College of Oceanography and Space Informatics, China University of Petroleum, Qingdao 266580, China (e-mail: heyw@upc.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2023.3261326

I. INTRODUCTION

OPTICAL satellite images are an important source of earth observation data, however, due to the inherent limitations of imaging system, optical remote sensing images are inevitably contaminated by clouds [1]. According to moderate resolution imaging spectrometer (MODIS) cloud mask product statistics, clouds cover approximately 67% of the earth's surface [2]. Cloud coverage in satellite images causes the loss of surface information, which not only limits data utilization but can also generate errors in the remote sensing application, such as land cover classification [3], surface temperature retrieval [4], and atmospheric variables estimation [5]. Therefore, rapid and high-precision cloud detection methods are critical for the fine processing and application of remote sensing imageries.

Over the past few decades, numerous cloud detection algorithms have been developed. Current cloud detection approaches are classified into three types: threshold-based, multitemporal-based, and machine learning methods. These methods use spectral, spatial, and temporal information, as well as joint features to separate clouds from the clear-sky areas. Threshold-based methods, such as the International Satellite Cloud Climatology Project algorithm and the Advanced Very High Resolution Radiometer processing scheme over the cloud land and ocean (APOLLO) algorithm [6], [7], use thresholds to identify clouds and underlying surfaces based on spectral differences or low brightness temperature. The MODIS cloud mask algorithm fully employs the rich spectral information of MODIS data and sets thresholds to generate cloud mask products [8], [9]. The function of mask (Fmask) algorithm proposed by Zhu and Woodcock [10] uses a series of thresholds to obtain cloud detection probability map for Landsat images. Fmask 4.0 [11] compensates for the lack of cloud and shadow detection in Fmask 3.3 and enhances the detection accuracy. It is challenging to determine a suitable threshold for large-scale remote sensing scenarios due to the diversity of surface types, particularly for thin clouds and highly heterogeneous surfaces such as desert, ice/snow, and bare rocks. To reduce the influence of fixed thresholds and complex surfaces, Sun et al. [12] proposed a dynamic threshold cloud detection method based on prior surface reflectance, which increased cloud detection accuracy over various underlying surfaces. Because the surface reflectance database is built using high-temporal-resolution images, their approach is best suited for images acquired in short revisit intervals. In general,

parameter adaptation and global optimization are typically difficult to perform for threshold-based methods due to the complexity of the surface environment and the diversity of cloud geometry, leading to varying degrees of cloud cover estimation bias [13].

Clouds are dynamic, and images obtained at different times in the same area can differ dramatically. Compared to the abrupt increase in surface reflectance generated by clouds, temporal changes in underlying surfaces are relatively smooth [14], [15]. Therefore, multitemporal-based methods are developed for cloud detection. Multitemporal information can alleviate the confusion of clouds and bright features. However, the implementation of this method necessitates the selection of cloud-free images or the construction of cloud-free reference images using multitemporal images [16], [17]. A clear-sky reference image is essential for cloud detection using multitemporal images. Obtaining cloud-free images is time-consuming and labor-intensive due to the effects of climate and sensor observing conditions, and cloud-free images in a particular location at a certain time are impossible to produce.

Machine learning-based approaches treat cloud detection as a binary (cloud and noncloud) or multiclass (thin cloud, thick cloud, and noncloud) task by establishing a reliable classifier and iteratively optimizing the parameters using massive training data [18]. Support vector machines [19], [20], random forests [21], [22], decision trees [23], and neural networks [24], [25], have been widely used in cloud detection. Traditional machine learning methods can provide representative features for identifying clouds and underlying surfaces; however, parameter selection, manual empirical judgment, and feature extraction are still required, and their performance is limited by classification frameworks, network structures, and capabilities [26]. As an extension of machine learning, deep learning fully excavates different scales of information using deep convolutional neural networks (DCNNs), and with robust feature representation capabilities, has produced promising results in cloud detection research [27], [28], [29], [30]. Initially, DCNN-based cloud detection methods used image patches to classify clouds and surfaces [31], [32], [33]. It enhances the accuracy and applicability of cloud detection compared with existing approaches. However, the image patch-based detection method has a local receptive field that ignores image neighborhood information. Therefore, a fully convolutional network (FCN) [34] is applied for cloud detection. FCN-based cloud detection is a pixel-level semantic-segmentation process. The UNet network [35], which fuses features between different layers and comprehensively extracts image information, is commonly used in cloud detection [29], [36], [37]. Jeppesen et al. [36] proposed RS-Net for achieving promising cloud detection results, although it had limited multispectral capabilities. To reduce the annotation of training samples, Ma et al. [13] combined ASTER library and AVIRIS spectral images using convolutional neural network (CNN) to achieve cloud detection for multi-sensor remote sensing imageries. Domain adaption [38], [39] were also introduced to improve the cloud detection performance. CNN-based cloud detection can fully mine the spectral and spatial information of images, and produce accurate cloud detection results,

however, it yields weaker performance in capturing global context information. Attention mechanism [40], [41] focusing on the important features was combined with CNN structure to improve cloud detection. Zhang et al. [42] introduced the spatial-channel attention mechanism into the cloud detection network, strengthened the feature information, and obtained better cloud detection results. Li et al. [43] proposed the Global Context Dense Blocks (GCDB-UNet) for cloud detection. This method embeds GCDB into the UNet framework and can detect thin clouds effectively.

MODIS images have a wide spectral range and high temporal resolution, and MODIS data are readily available; therefore, they are widely used in land, ocean, and atmosphere studies. An accurate cloud mask is crucial for the fine processing and application of MODIS images. At present, MODIS image cloud detection still has two aspects to be concerned. On the one hand, the effects of spatial resolution and mixed pixels (a mixture of the underlying surface and clouds) limit the cloud detection accuracy of MODIS. Deep learning technology can exploit the difference in spectral, spatial, and temporal information between clouds and the surface, indicating a significant application potential in MODIS image cloud detection. On the other hand, thin clouds over the inhomogeneous underlying surface and highly mixed scenes are the primary challenges in cloud detection. Changes in cloud radiation information across the land surface present complex uncertainties due to complex land surfaces and variable cloud phase states, sizes, and densities, particularly when pixels are covered by thin clouds. Satellite sensors provide relatively limited information, making it difficult to distinguish between highly heterogeneous surfaces and polymorphic clouds. The lack of land surface information hinders the distinction between clouds and complex surfaces. The combination of prior surface reflectance data and deep learning might be a promising method for MODIS image cloud detection.

To address the aforementioned issues, *a priori* land surface reflectance (LSR) dataset was coupled with a multiscale feature convolution neural network for MODIS imagery cloud detection (SRMF-CD). The LSR dataset constructed by MOD09 was employed to provide specific information for creating differential features with MODIS images. A multiscale feature cloud detection network using an atrous spatial pyramid pooling (ASPP) module and a channel and spatial attention module (CSAM) could integrate low-level spatial features and high-level semantic information to accurately separate clouds from surfaces.

The main contributions of this article can be summarized as follows.

1) A new cloud detection technique was developed using an LSR dataset and multiscale feature cloud detection network. The LSR dataset provides specific ground information without cloud coverage, hence solving the problem of obtaining a clear sky reference.

2) Difference-based training samples obtained using LSR dataset and MODIS images of different phases were used as the information source of the cloud detection network. Compared with the top of atmosphere (TOA) reflectance images, the differential images enhance the cloud feature and weaken the bright ground feature, alleviating the

TABLE I
BAND PARAMETERS AND THEORETICAL ACCURACY OF MOD09A1

Band	Spectral range	Signal to noise ratio	Absolute error	Relative error %
1	0.620–0.670	128	0.005	10–33
2	0.841–0.876	201	0.014	3–6
3	0.459–0.479	243	0.008	50–80
4	0.545–0.565	228	0.005	5–12
5	1.230–1.250	74	0.012	3–7
6	1.628–1.652	275	0.006	2–8
7	2.105–2.155	110	0.003	2–8

phenomenon of thin cloud omission and bright features misjudgment.

3) A multiscale feature cloud detection network (MFCD-Net) was designed for cloud detection. The DCNN, ASPP module, and attention module combined low-level spatial features and high-level features to capture multiscale information and generate a high-precision cloud mask.

II. DATA SOURCES

In this article, we used Terra MODIS images and MOD09 surface reflectance products to provide cloud and surface information (MODIS images and products are available at https://ladsweb.modaps.eosdis.nasa.gov/) [44]. MODIS is a medium-resolution imaging spectrometer that is carried by both the Terra and Aqua satellites, and it performs a complete scan of the earth's surface every one to two days. MODIS has 36 spectral bands, encompassing visible and infrared spectral ranges (0.4–14.4 μm) as well as image spatial resolution ranges from 250–1000 m. Therefore, it can provide large-scale global data such as cloud cover, radiance energy, and ocean and land changes.

The MOD09 dataset is a Level 2 product of the MOD09 surface reflectance series. It is a high-accuracy estimate of the surface spectral reflectance of each spectral channel, corrected for atmospheric conditions, aerosol scattering, thin clouds, cirrus clouds, and other factors [45]. MOD09A1 is an 8-day gridded Level 3 surface reflectance product with seven bands covering visible to near-infrared wavelengths. It delivers the optimal observations within eight days, effectively reducing the effect of surface and cloud interference. Under favorable conditions, the atmospheric correction accuracy was $\pm (0.005 + 0.05 \times \rho)$ [44], [46]. Table I lists the spectral range parameters of MOD09A1 as well as the theoretical errors resulting from atmospheric correction. The absolute error of each band of the MOD09A1 data was lower than 0.02, indicating that the data can accurately represent the actual surface reflectance.

III. METHODOLOGY

The surface reflectance of most objects was assumed to remain almost constant over a certain timeframe [12], [45]. Compared with the abrupt increase in surface reflectance generated by clouds, changes in the underlying surfaces were considered to be relatively smooth. This implied that the cloud-covered region differed significantly from the surface reference. The TOA reflectance curves of clouds and typical ground features are shown in Fig. 1, as are the reflectance difference curves between the TOA reflectance and the surface reference. As shown in Fig. 1(a), the TOA reflectance of snow is similar to that of thick clouds, which have the highest TOA reflectance, whereas the TOA reflectance of desert and bright artificial surfaces is similar to that of thin clouds. Due to the synergy of prior surface information, the reflectance difference in cloudy areas is most apparent in Fig. 1(b), and the reflectance differences in water, vegetation, desert, and snow are minor. Therefore, the reflectance difference information could effectively distinguish the cloud from the surface to a certain extent and reduce the confusion between the highlighted surfaces and clouds.

Here, the proposed algorithm consists of three parts: construction of the LSR dataset, generation of difference-based training samples, and cloud detection using a multiscale features network. The framework of the proposed algorithm is shown in Fig. 2. The LSR dataset constructed using MOD09A1 provided the real surface information as clear-sky references. Difference-based training samples were generated using difference operations. The multiscale feature cloud detection network utilized low-level spatial features, high-level semantic features, and temporal information to achieve high-precision cloud detection.

A. Construction of LSR Dataset

To construct a high-quality LSR dataset, MOD09A1 surface reflectance products from 2014 to 2018 were downloaded in this study. The visible and near-infrared bands are effective in providing spectral information and are common channels in most sensors. Considering the universality of the proposed method in cloud detection of different types of satellite images, bands 1, 2, 3, and 4 of the MODIS images are used in this article. To ensure the spatial continuity of surface reflectance data and the accuracy of the spectral information, the LSR database was constructed using the monthly minimum synthesis method [12], as follows:

$$L(i,j) = \mathrm{Min}((L_1(i,j), L_2(i,j), L_3(i,j), L_4(i,j)) \quad (1)$$

where L is the synthetic LSR image, L_1, L_2, L_3, and L_4 are the MOD09A1 data of four scenes in a month, and i, j are the rows and columns of a scene image.

According to (1), a LSR dataset consisting of 12 global monthly composite surface reflectance images in the visible and near-infrared bands (blue, green, red, and near-infrared) was constructed. The projection and coordinate systems of the LSR dataset were unified into the Albers projection and WGS84 coordinate systems, respectively. The constructed dataset could reduce the influence of clouds and cloud shadows and consider
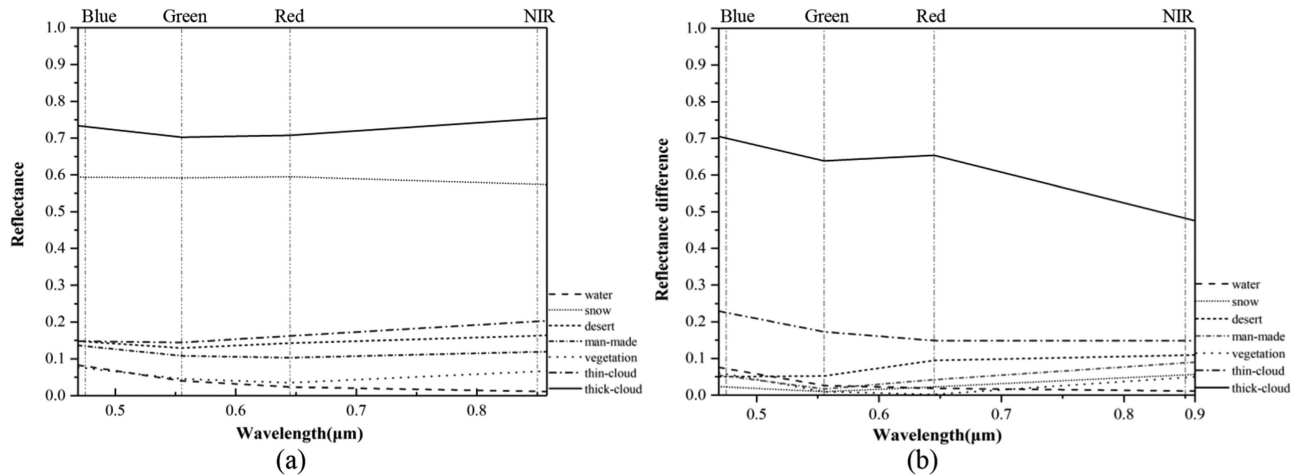
Fig. 1.   Reflectance curves of cloud and typical objects from MODIS images. (a) TOA reflectance. (b) Reflectance difference.
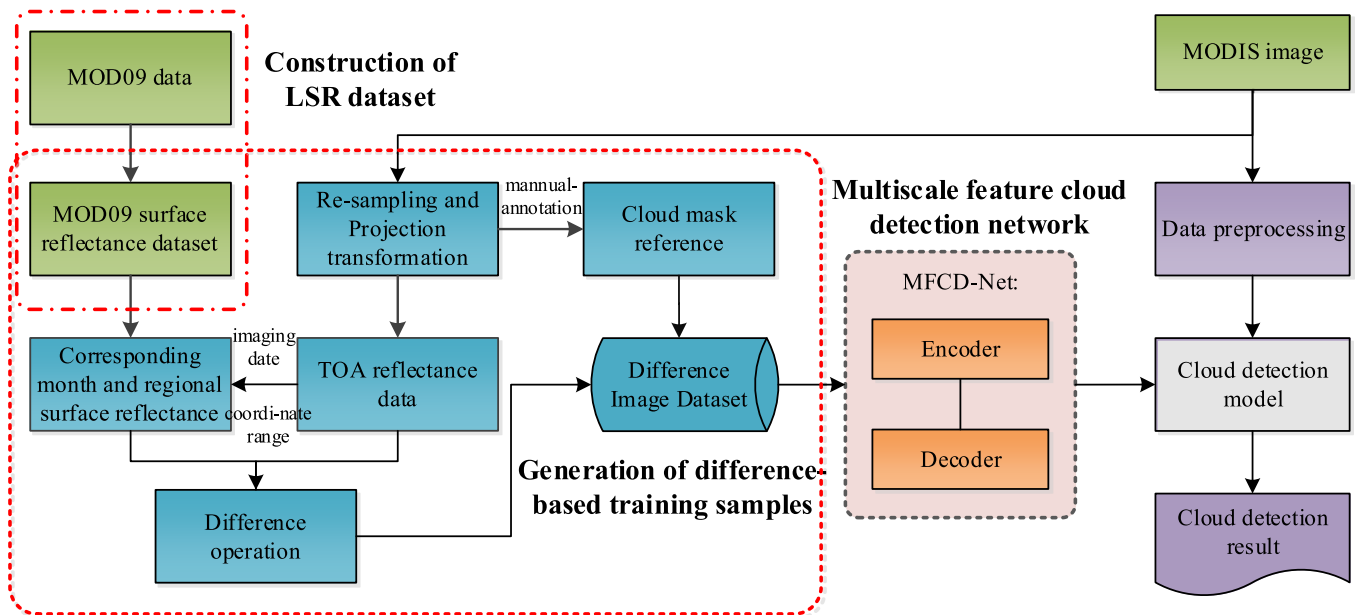

Fig. 2.   Flowchart of the proposed algorithm based on land surface reflectance and the multiscale feature cloud detection network (SRMF-CD).

the temporal changes of the surface, enabling it to effectively adjust to seasonal changes while maintaining strong spatial continuity. Fig. 3 shows a false-color synthesis image of global surface reflectance in June 2014, demonstrating strong spatial continuity and accurate representation of the real surface, with slight effects from factors such as clouds and cloud shadows.

B. Generation of Difference-Based Training Samples

The differential image was generated using the constructed surface reflectance image and the MODIS image to be detected. First, the MODIS image was radiometrically calibrated to obtain TOA reflectance. Subsequently, the surface reflectance image of the corresponding month and location was obtained according to the time, longitude, and latitude of the cloud coverage image.

Finally, the surface reference image and the TOA reflectance image of the same area and month were used for the difference calculation to obtain the differential image. The equation is as follows:

$$I_{Dk}(i,j) = |I_{Tk}(i,j) - I_{Sk}(i,j)| \tag{2}$$

where I_{Dk} represents the difference image, I_{Tk} represents the TOA reflectance image, I_{Sk} represents the surface reference image, the k represents the band, where k is 1, 2, 3, and 4, and i and j represent the ith row and jth column, respectively.

In this article, 61 MODIS images with different spatial distributions at different times including distinct underlying surface types, such as vegetation, water bodies, urban regions, deserts, and ice/snow, were employed, of which 45 scenes were used for training and 16 scenes for validation. The cloud masks were
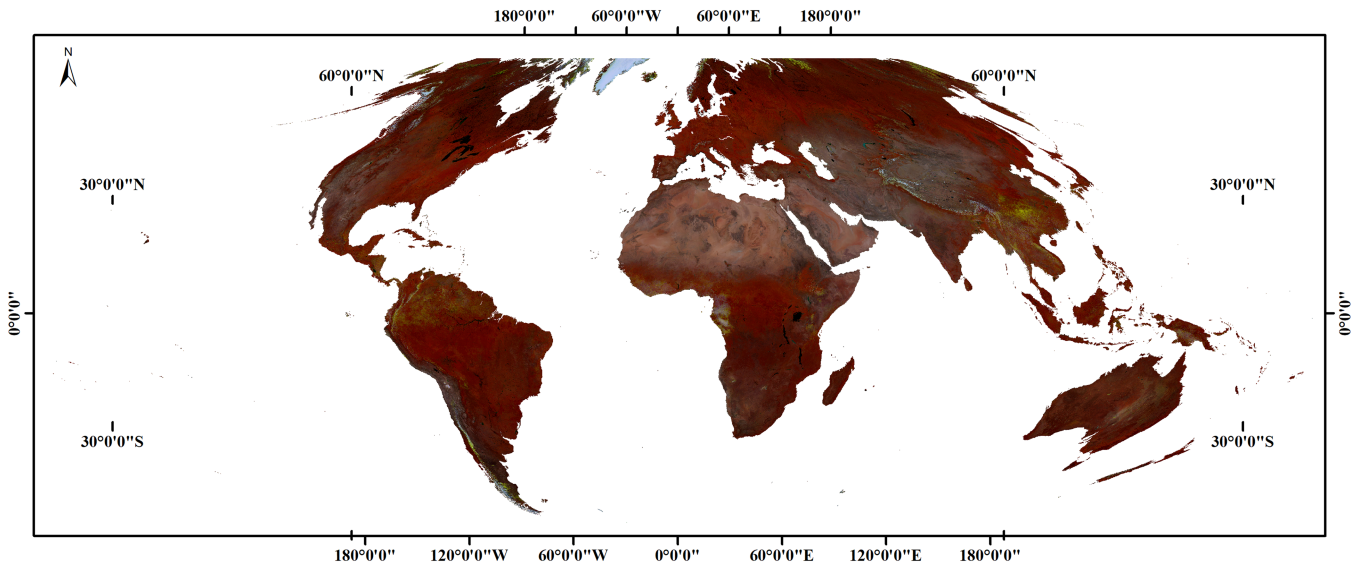
Fig. 3. False color synthesis image (RGB: bands 2, 4, 3) of global surface reflectance in June 2014.

manually annotated, where pixel value 1 indicates cloud and pixel value 0 indicates no cloud. Considering the limitations of the network structure and hardware facilities, the entire scene of the differential images and cloud masks were cropped into images with a size of 512×512 pixels. Ultimately, the difference-based training samples consist of more than 4200 training images.

Table II shows the MODIS true-color images and the corresponding differential images for various surface types. The first and third columns display the true-color composite images (RGB: bands 1, 4, and 3), and the second and fourth columns display the differential images. As shown in Table II (blue arrow), bright surface information in the differential image, such as rocks, deserts, and ice/snow are weakened, while clouds are highlighted, which is important for distinguishing clouds from highlighted surfaces. Additionally, the differential image enhances the thin cloud features.

C. Cloud Detection Using Multiscale Feature Network

Due to the phenomenon of same spectrum foreign matter and the limited number of extracted spectral features, distinguishing clouds from complex surfaces based solely on spectral features is difficult. However, DCNNs have achieved remarkable performance by combining spatial and spectral information, thanks to their powerful information extraction and feature representation capabilities. Thus, a multiscale feature cloud detection network was designed in this article using the differential feature dataset as the primary source of information.

Network Structure: The cloud detection network adopts an encoder–decoder structure, as shown in Fig. 4. In the encoder section, a DCNN uses a differential image as input to obtain low-level and high-level feature. The high-level feature is fed into an ASPP module to obtain multiscale semantic information. However, the convolution operation at different scales may cause information loss. To enhance the feature representation ability,
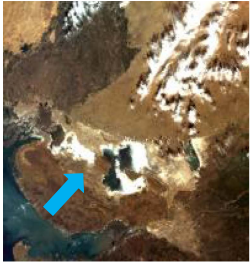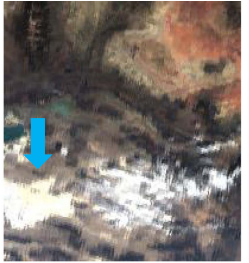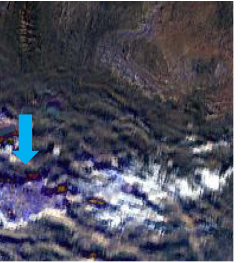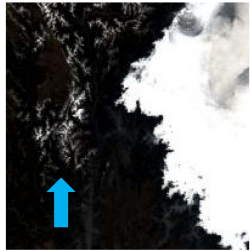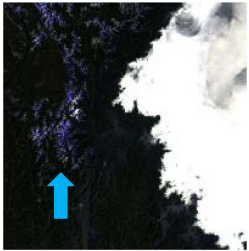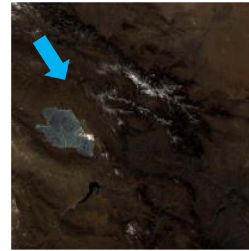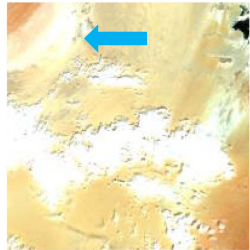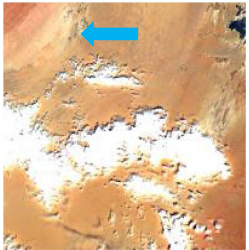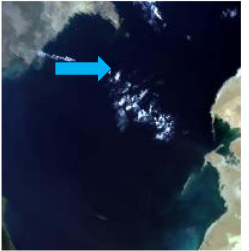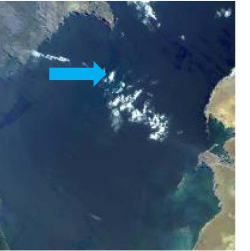
a CSAM is designed to learn high-level feature, as shown in Fig. 5. First, max-pooling and average pooling operations extract various channelwise attention features. Then, shared fully connected layers learn channel correlations and weight distributions. After fusing the two outputs, an attention map is obtained by applying a sigmoid function. The re-weighted feature maps are generated by multiplication operations bet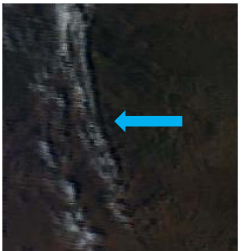ween the attention map and F_h. Subsequently, the channel-refined feature maps are fed into the spatial attention mechanism, which differs from the channel attention mechanism in that the 1×1 convolutional layer replaced the shared fully connected layer. Finally, the attention feature map F_A is generated. A concatenation operation is used to merge the two types of feature maps from the CSAM module and ASPP modules.

The decoder first performed a four-fold bilinear upsampling on the multiscale semantic feature and then combined the low-level features to refine the concatenated features using 3×3 convolutions. The final pixel-level cloud mask was generated via four-fold bilinear upsampling. The cloud detection network combined the respective advantages of the ASPP module and the CSAM module to produce clearer segmentation object boundaries while effectively capturing multiscale features.

Network Training Details: The training of the cloud detection model is the process of continuously optimizing the parameters to achieve pixel-level recognition of clouds and the ground surface. The difference-based samples with red, blue, green, and near-infrared bands were used for input images. The total number of training images is 4267, and the ratio of validation images to training images is 1:9. To minimize the discrepancy between predictions and ground truth, the adaptive moment estimation (Adam) optimizer was used to dynamically optimize the model parameters. The initial learning rate was set to 0.001 and the batch size was set to four.

Furthermore, to address the issue of sample imbalance, a novel loss function that combines the Dice loss and weighted binary cross-entropy (DWCE) loss has been introduced. This function

TABLE II
MODIS IMAGE AND CORRESPONDING DIFFERENCE IMAGE OVER VARIOUS TYPES OF UNDERLYING SURFACES

MODIS image	Difference image	MODIS image	Difference image
Bare soil and rocks		Bare soil and rocks	
Snow/ice		Snow/ice and bare rocks	
Desert		Water	
Bare soil and rocks		Vegetation	

aims to mitigate the influence of imbalanced data on the training process. The specific definition of this function is as follows:

$$L_{\text{DWCE}} = L_{\text{Dice}} + L_{\text{WCE}} \tag{3}$$

where L_{WCE} represents WCE loss, and L_{Dice} represents soft Dice loss. The WCE loss gives weights to different classes to reduce model bias due to sample imbalance (the number of noncloud pixels is greater than that of cloud pixels). It is defined as follows:

$$L_{\text{WCE}} = -\sum_{c=1}^{M} w_c y_c \log(p_c)$$

$$w_c = \frac{N - N_c}{N} \tag{4}$$

where w_c is the weight of each category, N is the total number of pixels, and N_c denotes the number of pixels of the true category
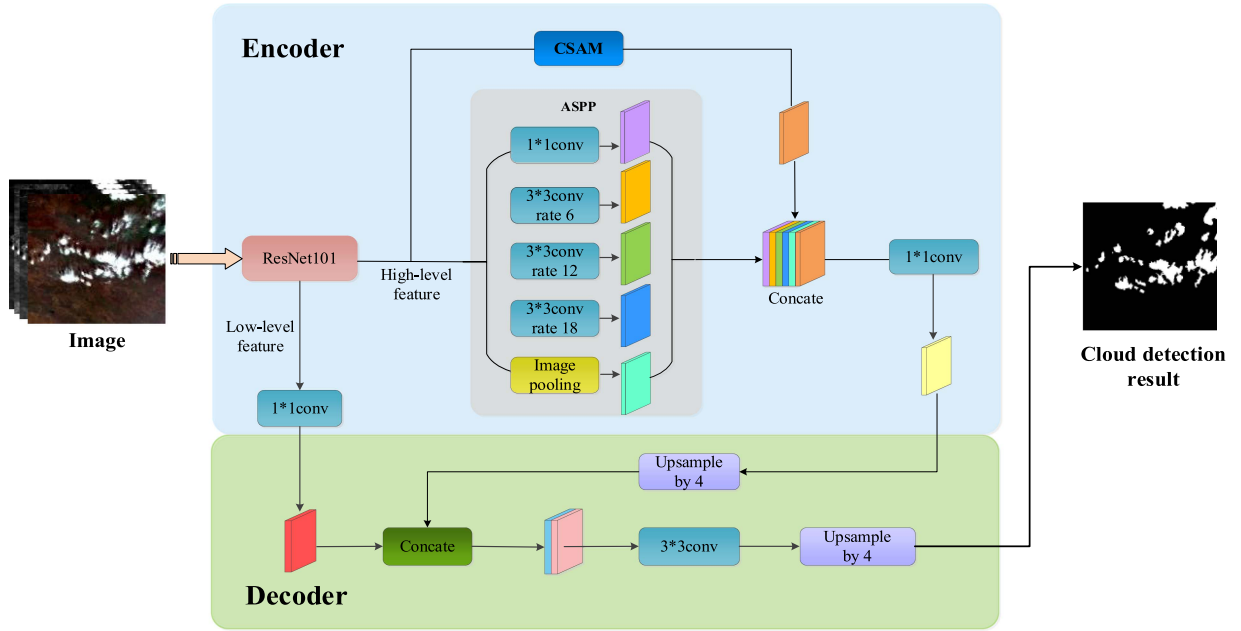
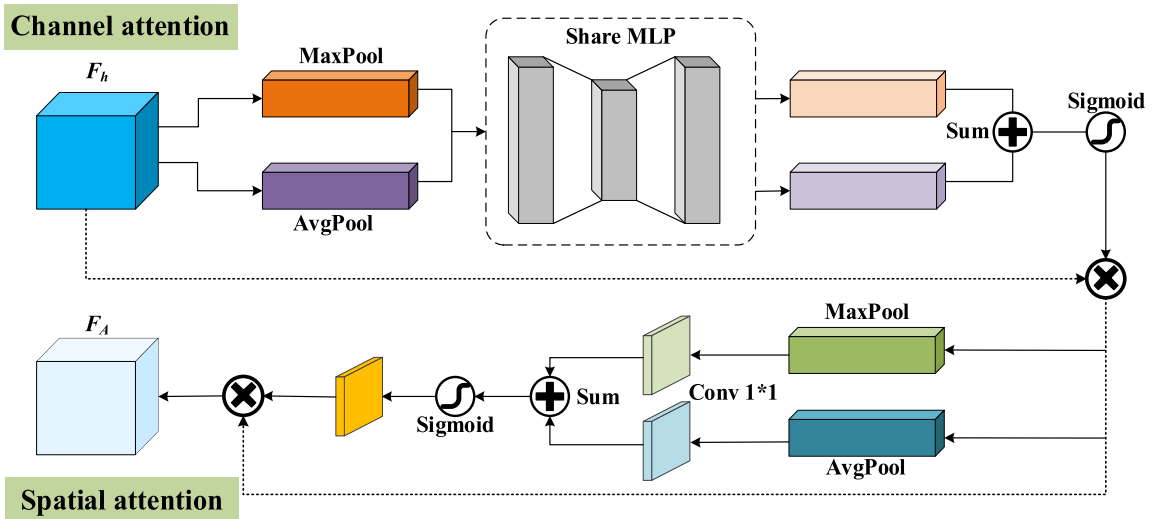Fig. 4.   Flowchart of the multiscale feature cloud detection network.

Fig. 5.   CSAM module. The upper part is the channel attention mechanism and the lower part is the spatial attention mechanism.

c. Dice loss is defined as follows:

$$L_{\text{Dice}} = 1 - \frac{2TP}{2TP + FN + FP} \qquad (5)$$

where TP denotes the cloud pixels, FP denotes the noncloud pixels detected as cloud pixels, and FN denotes the noncloud pixels detected as cloud pixels. The DWCE loss takes the advantages of WCE loss and Dice loss to get a more stable convergence of the model.

The cloud detection model was implemented using Python 3.7 and PyTorch 1.7. The DCNN, ASPP module, and CSAM module provide multiscale features for cloud and ground surface pixel prediction. Based on training accuracy and loss, the optimal

cloud detection model was obtained when the number of epochs reached 100. Fig. 6 shows the feature map size of different layers.

IV. EXPERIMENT AND RESULTS ANALYSIS

A. Cloud Detection Results of MODIS Image

Sixteen MODIS images with an approximate cloud coverage of 5% to 100% were used to validate the performance of the SRMF-CD algorithm. Fig. 7 shows four scenes of cloud detection results over different underlying surfaces: vegetation, water, desert, and bare soil. White pixels indicate the cloud pixels and black pixels indicate the noncloud pixels. Compared with the MODIS true-color composite image, surface reflectance
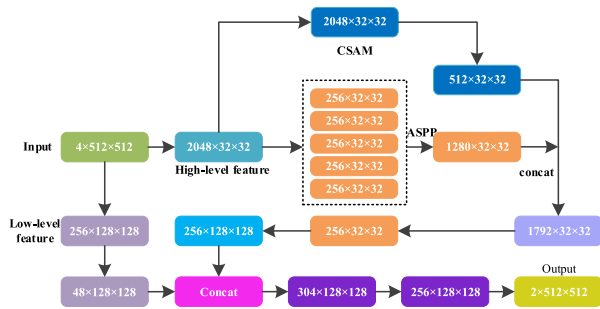
Fig. 6. Feature map size of different layers.

reference, and manual cloud mask, the proposed algorithm performed well on cloud detection in different scenes and achieved a complete cloud contour that is highly consistent with the manual cloud mask.

The detection of thin clouds presents a challenge for cloud detection algorithms, especially in heterogeneous regions where the surface reflectance varies greatly. Additionally, clouds may be misclassified due to the underlying surface type. For instance, in deserts, ice, snow, and other areas with high surface reflectivity, cloud pixels can be easily confused with the underlying surface. To evaluate the performance of the SRMF-CD algorithm in cloud detection for thin clouds and highly heterogeneous surfaces, we conducted a comparative experiment using a cloud detection method based on TOA reflectance data and MFCD-Net (MFCD-TOA). This method used the same network structure as the SRMF-CD algorithm but employed different inputs (TOA reflectance images).

Fig. 8 shows the cloud detection results of the SRMF-CD and MFCD-TOA methods three different surface types: 1) bright rocks, 2) ice/snow, and 3) water bodies. In Fig. 8(a)–(d), rocks and ice/snow with similar TOA reflectance to clouds appear white, and are often mistakenly identified as clouds in cloud detection based on spectral information. However, in the differential image, the reflectance changes of the highlighted rocks and ice/snow are small, which reduces confusion between rocks, ice/snow, and clouds. Therefore, the SRMF-CD algorithm improves the accuracy of cloud detection for bright surfaces owing to the prior surface information, while MFCD-TOA method incorrectly detects bright rocks and ice/snow as clouds. Because solar radiation of the thin cloud can easily penetrate the cloud layer to reach the ground, the reflectance characteristics of the thin cloud are similar to the reflectance characteristics of the underlying surface, and the difference in reflectivity between the two becomes smaller, making the detection of thin clouds another challenge in this field. In Fig. 8(e) and (f), the underlying surface type can be seen in the images covered by thin clouds, which has little impact on image interpretation. However, thin clouds are usually omitted in spectrum-based cloud detection, which can greatly impact applications such as parameter inversion and change monitoring. According to the difference image and spectral difference curve in Fig. 2, the reflectance difference between water bodies and vegetation surfaces is extremely small

(less than 0.1), while the difference in thin cloud areas is often greater than 0.1, especially in the blue band. The differential image enhances thin cloud features, promoting the multiscale feature-cloud detection network to extract effective information. The SRMF-CD method can better identify thin clouds with fewer omissions and has higher consistency with cloud masks, while MFCD-TOA method omits thin clouds over water, desert, and bare soil.

To further evaluate the effectiveness of the algorithm, we compared the cloud detection results of our proposed SRMF-CD method with those of UNet, Deeplabv3+, and MFCD-TOA methods. Fig. 9 presents the comparison of cloud detection results from different methods with cloud masks. The red circle indicates commission and the yellow rectangle indicates cloud omission. The UNet network obtains limited information due to the loss of information in the down-sampling layer, resulting in the omission of clouds. Although the Deeplabv3+ method improves the cloud detection results due to atrous convolution and feature integration, there are still more thin clouds missing, and the overall contours of the clouds are rather coarse. In contrast, the proposed multiscale feature cloud detection network that combines the ASPP module and the CSAM module is beneficial to extract multiscale information, which improves the accuracy of cloud recognition. Although the MFCD-TOA method has fewer clouds omitted, it misidentifies bright features as clouds. Based on the true-color image, surface reflectance image, and differential image, the change of the bright surface in the difference image was very small compared to the cloudy area. The SRMF-CD method, supported by surface reflectance, reduces the omission of thin clouds and the misidentification of the highlighted surface, and achieved the best cloud detection results among the four methods.

To fully demonstrate the role of the different components of the proposed method, we conducted ablation experiments for the differential feature, the attention mechanism module, and loss function, respectively. Fig. 10 presents the results of the ablation experiments and comparisons with state-of-the-art methods. As shown in Fig. 10(c1) and (h1), (c2) and (h2), the results using differential features as network inputs surpass those using TOA reflectance data as inputs, reducing the misidentification of bright surfaces. Fig. 10(d1) and (h1), (d2) and (h2) demonstrates that the inclusion of the attention mechanism module enhances cloud detection accuracy. Fig. 10(e1) and (h1), (e2) and (h2) suggests that the utilization of the DWCE loss function reduces cloud omission errors. Compared to advanced cloud detection methods such as UNet++ and PSPNet, the SRMF-CD algorithm generally outperforms these outstanding segmentation models in terms of accuracy. Specifically, it improves the detection of thin clouds and clouds in areas with high surface reflectance, resulting in cloud regions that are highly consistent with manual cloud masks.

B. Quantitative Assessment

To objectively evaluate the algorithm, the overall accuracy (OA), precision, recall, F1-Score, and Kappa were introduced
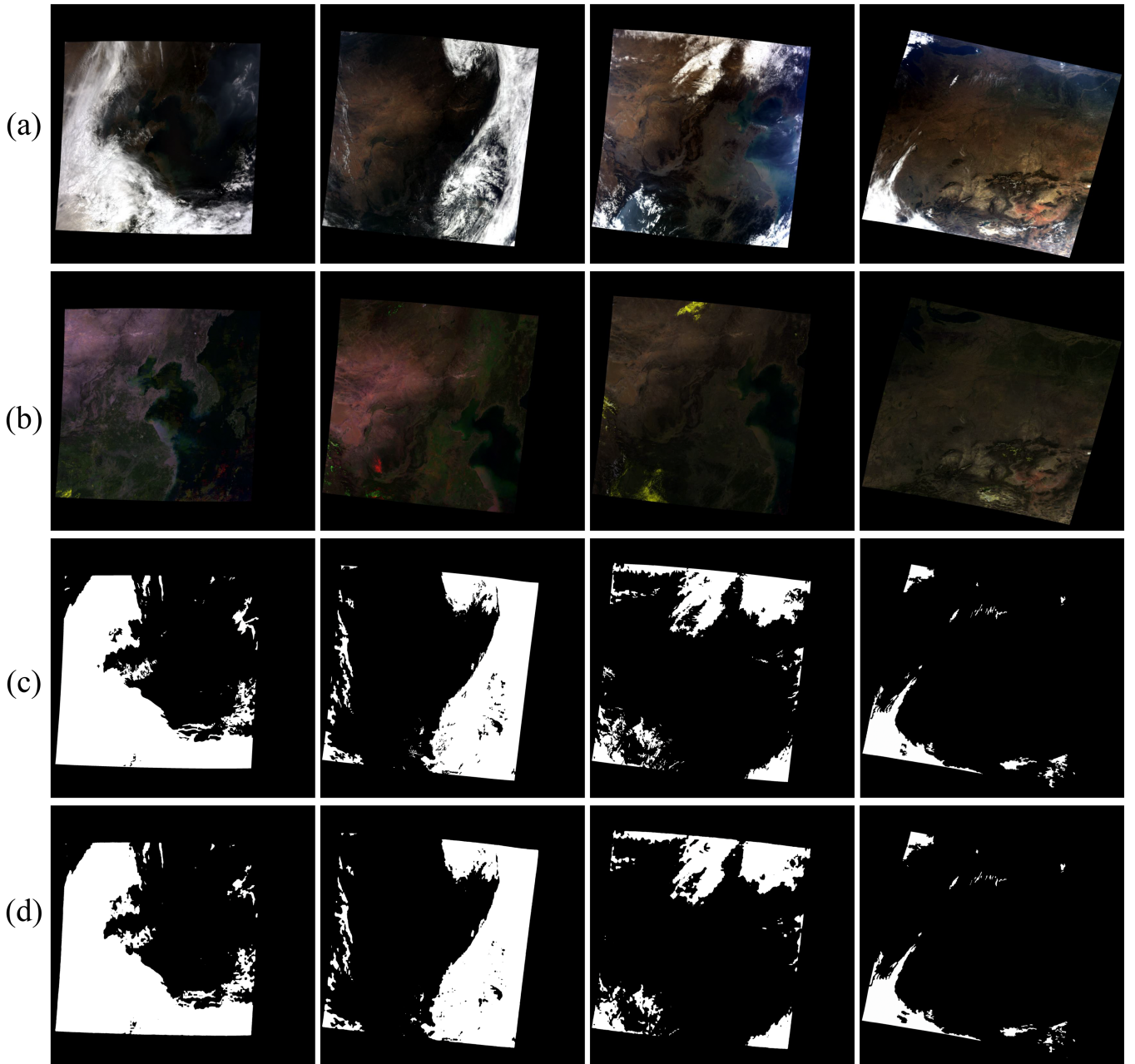
Fig. 7.   Cloud detection result of thick cloud and thin cloud over vegetation area, ocean area, urban area and desert area. (a) True-color image with red, green, and blue band. (b) Surface reflectance reference. (c) Manual cloud mask. (d) Cloud detection result of SRMF-CD method.

to quantitatively assess the cloud detection results of the different methods. These metrics are defined as

$$\text{OA} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \qquad (6)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad (7)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (8)$$

$$\text{F1} - \text{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (9)$$

$$\text{Kappa} = \frac{P_a - P_e}{1 - P_e} \qquad (10)$$

$$P_a = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \qquad (11)$$

$$P_e = \frac{(P^*(\text{TP} + \text{FP}) + N^*(\text{FN} + \text{TN}))}{(P+N)^2} \qquad (12)$$

where true positive (TP) denotes the total number of cloud pixels correctly predicted, true negative (TN) denotes the total number of non-cloud pixels correctly recognized, and false positive (FP) and false negative (FN) denote the total number of pixels with
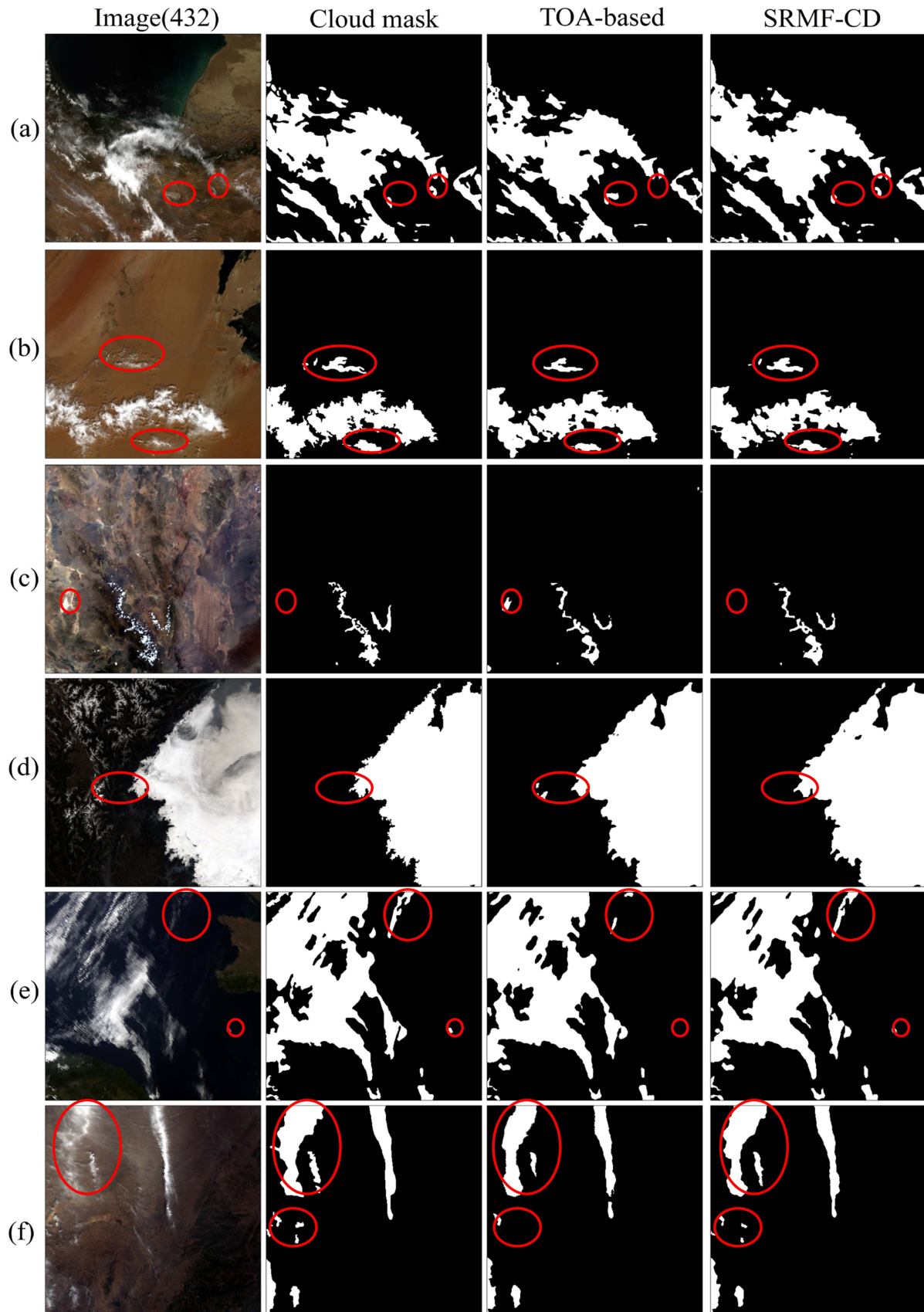
Fig. 8. Cloud detection results of different method over various underlying surface. (a) Water. (b) Desert. (c) Bare soil. (d) Ice/snow. (e) Water. (f) Vegetation.
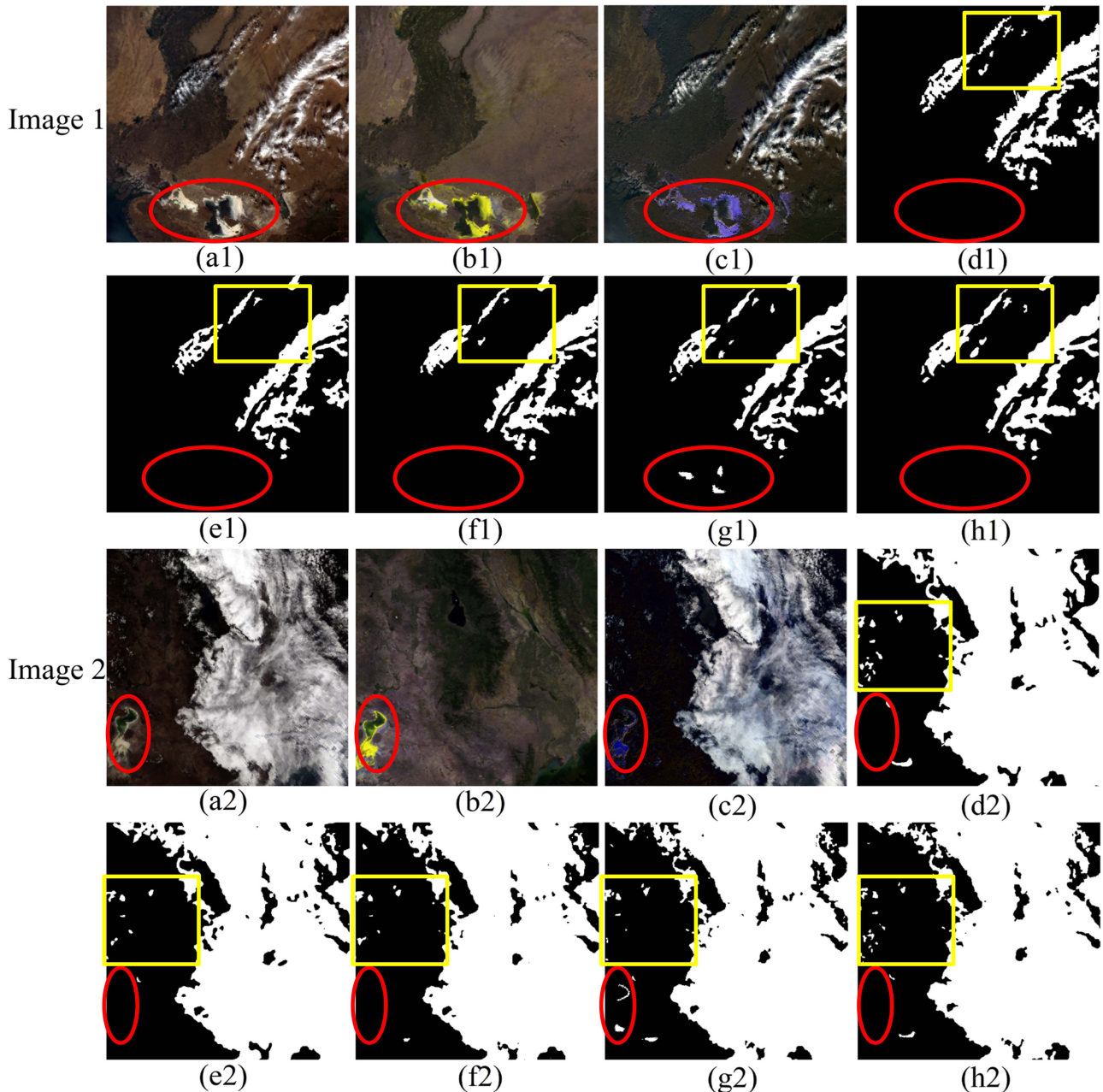
Fig. 9. Comparison of cloud detection results obtained by different methods. (a1) and (a2) True-color image with red, green, and blue band. (b1) and (b2) Surface reflectance reference. (c1) and (c2) Differential image. (d1) and (d2) Manual cloud mask. (e1) and (e2) Cloud detection results obtained by UNet method. (f1) and (f2) Cloud detection results obtained by Deeplabv3+ method. (g1) and (g2) Cloud detection results obtained by MFCD-TOA method. (h1) and (h2) Cloud detection results obtained by SRMF-CD method.

an incorrect outcome from cloud and noncloud recognition, respectively.

Table III presents the quantitative validation results of the UNet, Deeplabv3+, MFCD-TOA, UNet++, PSPNet, SRMF-CD-without-CSAM, SRMF-CD-without-DWCE, and SRMF-CD methods. The MFCD-TOA, SRMF-CD-without-CSAM, and SRMF-CD-without-DWCE methods represent ablation experiments. The MFCD-TOA method refers to the use of TOA reflectance data as the network input, the SRMF-CD-without-CSAM refers to the SRMF-CD method without the attention mechanism module, and the SRMF-CD-without-DWCE refers

to the SRMF-CD method without the use of the DWCE loss function.

As shown in Table III, the SRMF-CD method achieved outstanding cloud detection performance with high accuracy. The OA reached 96.55%, whereas the precision and recall reached 92.13% and 88.90%, respectively, indicating fewer omissions and commission errors of cloud pixels. The F1-score and Kappa were 90.44% and 88.85%, respectively, objectively demonstrating that the SRMF-CD algorithm exhibited good detection performance over diverse underlying surfaces. Furthermore, the cloud detection results of UNet, Deeplabv3+, UNet++, and
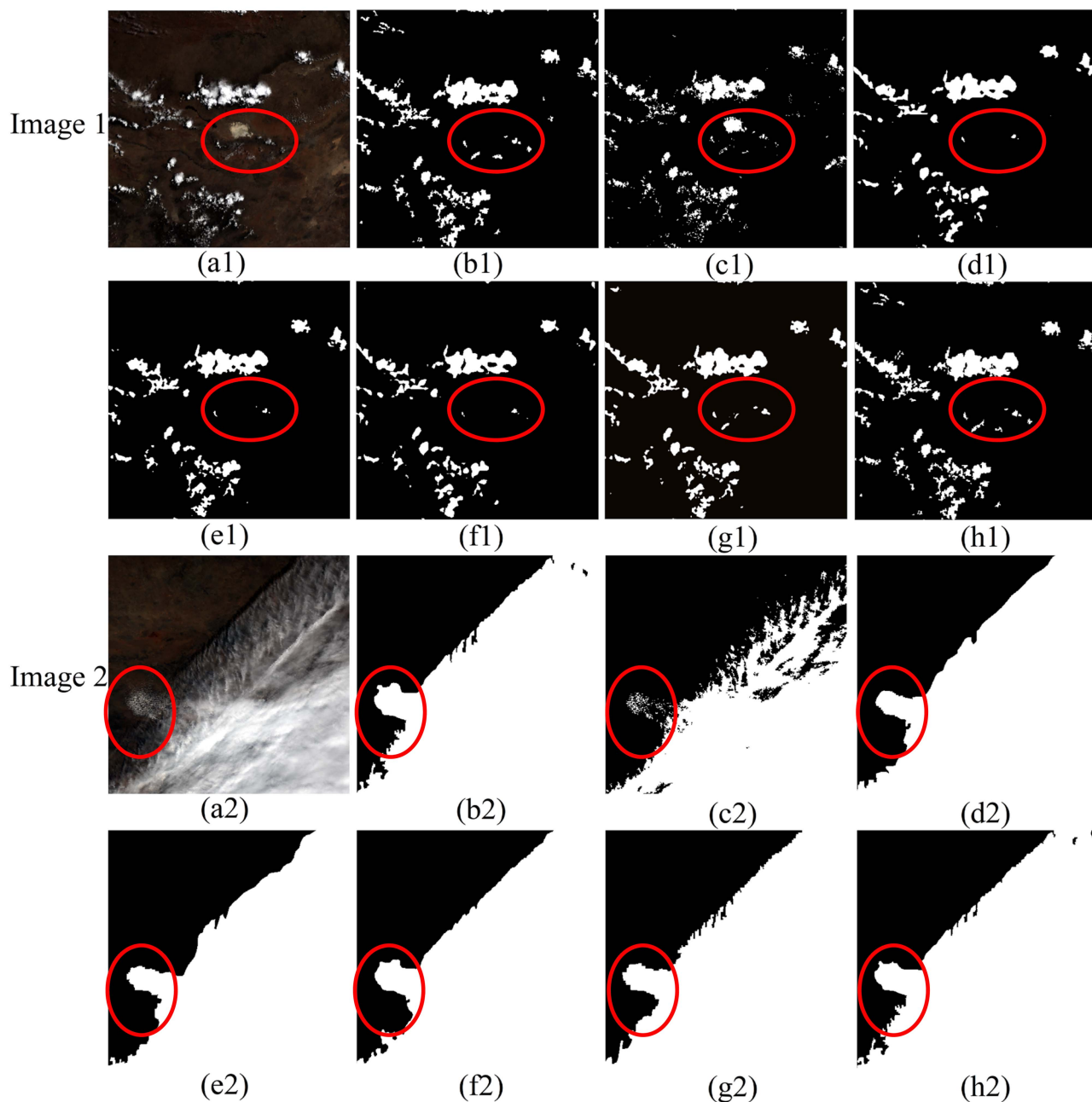
Fig. 10.   Results of the ablation experiments and the comparison with UNet++ and PSPNet methods. (a1) and (a2) True-color image with red, green, and blue band. (b1) and (b2) Manual cloud mask. (c1) and (c2) Cloud detection results with TOA reflectance as network input. (d1) and (d2) Cloud detection results without CSAM module. (e1) and (e2) Cloud detection results without DWCE loss. (f1) and (f2) Cloud detection results obtained by UNet++ method. (g1) and (g2) Cloud detection results obtained by PSPNet method. (h1) and (h2) Cloud detection results obtained by SRMF-CD method.

PSPNet methods are inferior to those of the proposed cloud detection network (MFCD-TOA and SRMF-CD), which is due to the fact that the multiscale feature cloud detection network provides richer spectral and spatial information. Benefiting from stable surface reflectance and cloud reflectance differences, the SRMF-CD methods improve the accuracy of cloud pixels and reduce the misclassification of cloud pixels and clear-sky pixels. Compared with the results of SRMF-CD-without-CSAM and SRMF-CD without-DWCE methods, the attention mechanism and DWCE loss improved the cloud detection accuracy.

C. Efficiency

All the experiments were conducted on a desktop computer with an Intel Core i5-9500 (3.10 GHz), an NVIDIA GEFORCE RTX 2080 Ti GPU (with 11 GB) and 32 G DDR4 Memory.

Training a model took approximately 16 hours, but predicting a 512×512 image only took 0.4 s. An entire MODIS scene can be predicted in less than 20 s. Compared to other methods, such as SRMF-CD-without-CSAM and Deeplabv3+, the SRMF-CD method did not add additional prediction time. This suggests that

TABLE III
QUANTITATIVE ANALYSIS OF DIFFERENT METHODS

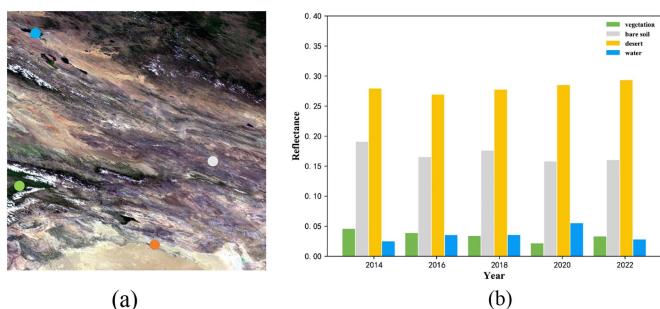| Method	OA (%)	Precision (%)	Recall (%)	F1-score (%)	Kappa (%)
UNet	94.64	91.85	71.45	79.54	76.87
Deeplabv3+	95.00	91.38	72.12	80.86	77.84
MFCD-TOA	96.12	92.06	84.72	88.12	86.27
UNet++	95.57	91.92	83.25	85.16	84.33
PSPNet	95.89	92.01	85.02	88.04	86.19
SRMF-CD-without-CSAM	95.60	91.53	80.75	83.37	82.91
SRMF-CD without-DWCE	95.83	91.69	84.18	86.58	84.60
SRMF-CD	96.55	92.13	88.90	90.44	88.85

Fig. 11. Analysis of surface reflectance of different underlying surfaces from 2014 to 2022. (a) True color composite image, location at 92°46′43″E, 45°24′30″N. (b) Histograms of reflectance over different underlying surface in 2014, 2016, 2018, 2020, and 2022 year.

processing MODIS images at a rapid pace is achievable using a standard desktop computer equipped with a single GPU.

D. Disscussion

Qualitative and quantitative evaluations have shown that the SRMF-CD algorithm performs well in MODIS image cloud detection. To improve thin cloud detection and reduce misclassification between bright surface and cloud, we synergized the prior LSR dataset with a multiscale feature convolutional neural network, providing a promising strategy for cloud detection. The algorithm is developed based on the assumption that the land surface reflectance changes little within a certain period. To evaluate the change of the surface reflectance, we studied the change of the surface reflectance of the vegetation, water, desert, and bare soil areas from 2014 to 2022 year. The MOD09A1 product in the same area in 2014, 2016, 2018, 2020, and 2022 year were selected to analyze the reflectance differences of different underlying surface. Fig. 11 shows the MODIS image and histograms of reflectance over different underlying surface. In Fig. 11(a), the dot indicates the selected underlying surface location. From Fig. 11(b), it can be seen that the surface reflectance of vegetation, water bodies, deserts, and bare soil are all relatively stable from 2014 to 2022, with differences well below 0.1. This indicates that the surface reflectance of most objects changes little over several years. Therefore, the surface reflectance image as a reference can reflect the real surface information, and the difference feature can effectively highlight the existence of clouds.

Although the SRMF-CD algorithm has good adaptability to most surface environments, it may produce weak performance in some areas where the surface reflectance changes obviously, such as snowfall/melt, natural disasters, urban sprawl, and so on.

V. CONCLUSION

This article proposes a novel cloud detection method based on a land surface reflectance dataset and multiscale feature network to improve cloud detection for MODIS images. To provide prior land surface information, the global monthly synthetic LSR dataset constructed by MOD09 products is used as land surface reference. Difference-based training samples, including red, green, blue, and near-infrared bands, are generated by the difference operation of surface reflectance image and the MODIS image. A cloud detection network coupled with an ASPP module and a CSAM module is designed to identify clouds and ground surfaces by mining differential image information and multiscale feature representation. The findings show that the SRMF-CD method performes well in cloud detection in different conditions.

Furthermore, compared with the qualitative and quantitative analyses based on UNet, Deeplabv3+, UNet++, PSPNet, and MFCD-TOA methods, the proposed method obtains the highest OA (96.55%), precision (92.13%), recall (88.90%), F1-score (90.44%), and Kappa (88.85%), demonstrating that the SRMF-CD algorithm alleviates cloud pixel omissions and commission errors. The SRMF-CD method, which benefits from the weakening of the bright surface and cloud information enhancement in difference images, can effectively reduce the effects of mixed pixels and achieve cloud detection from a large-scale area and long-term sequence.

With the assumption that the surface reflectance of most features varies little over time, the algorithm has certain limitations in some areas where the surface reflectance changes obviously, due to snowfall/melt, natural disasters, urban sprawl, and so on. As new satellites are deployed and the source of satellite data grows, and the generality of the algorithm is also worth

considering. In future work, we will investigate the effects of major changes in surface reflectance and develop novel cloud detection methods for various satellite data.

REFERENCES

[1] Z. Li, H. Shen, Q. Weng, Y. Zhang, P. Dou, and L. Zhang, "Cloud and cloud shadow detection for optical satellite imagery: Features, algorithms, validation, and prospects," *ISPRS J. Photogrammetry Remote Sens.*, vol. 188, pp. 89–108, Jun. 2022.

[2] Y. Zhang, W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko, "Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data," *J. Geophys. Res., Atmos.*, vol. 109, Oct. 2004, Art. no. D19105, doi: 10.1029/2003JD004457.

[3] J. Wei et al., "Cloud detection for landsat imagery by combining the random forest and superpixels extracted via energy-driven sampling segmentation approaches," *Remote Sens. Environ.*, vol. 248, Oct. 2020, Art. no. 112005.

[4] H. Ma, S. Liang, H. Shi, and Y. Zhang, "An optimization approach for estimating multiple land surface and atmospheric variables from the geostationary advanced himawari imager top-of-atmosphere observations," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 2888–2908, Apr. 2021, doi: 10.1109/TGRS.2020.3007118.

[5] J. Wei et al., "An improved highspatial-resolution aerosol retrieval algorithm for MODIS images over land," *J. Geophys. Res. Atmos.*, vol. 123, no. 21, pp. 12,291–12,307, Oct. 2018, doi: 10.1029/2017JD027795.

[6] W. Rossow and E. Dueñas, "The international satellite cloud climatology project (ISCCP)," *Bull. Amer. Meteorological Soc.*, vol. 85, no. 2, pp. 167–172, 2004.

[7] R. Saunders and K. Kriebel, "An improved method for detecting clear sky and cloudy radiances from AVHRR data," *Int. J. Remote Sens.*, vol. 9, no. 1, pp. 123–150, 1988, doi: 10.1080/01431168808954841.

[8] S. A. Ackerman, K. I. Strabala, and W. P. Menzel, "Discriminating clear sky from clouds with MODIS," *J. Geophys. Res.*, vol. 103, no. D24, pp. 32141–32157, Dec. 1998.

[9] R. A. Frey, S. A. Ackerman, and Y. Liu, "Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5," *J. Atmospheric Ocean. Technol.*, vol. 25, no. 7, pp. 1057–1072, Jul. 2008.

[10] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadowdetection in Landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83–94, Mar. 2012.

[11] S. Qiu, Z. Zhu, and B. He, "Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4-8 and Sentinel-2 imagery," *Remote Sens. Environ.*, vol. 231, Sep. 2019, Art. no. 111205.

[12] L. Sun et al., "A universal dynamic threshold cloud detection algorithm (UDTCDA) supported by a prior surface reflectance database," *J. Geophys. Res. Atmos.*, vol. 121, no. 12, pp. 7172–7196, Jun. 2016.

[13] N. Ma, L. Sun, C. Zhou, and Y. He, "Cloud detection algorithm for multi-satellite remote sensing imagery based on a spectral library and 1D convolutional neural network," *Remote Sens.*, vol. 13, no. 16, Aug. 2021, Art. no. 3319, doi: 10.3390/rs13163319.

[14] T. Bai, D. Li, K. Sun, Y. Chen, and W. Li, "Cloud detection for high-resolution satellite imagery using machine learning and multi-feature fusion," *Remote Sens.*, vol. 8, no. 9, Aug. 2016, Art. no. 715, doi: 10.3390/rs8090715.

[15] R. Liu and Y. Liu, "Generation of new cloud masks from MODIS land surface reflectance products," *Remote Sens. Environ.*, vol. 133, pp. 21–37, Jun. 2013, doi: 10.1016/j.rse.2013.01.019.

[16] L. Gomez-Chova et al., "Cloud masking and removal in remote sensing image time series," *J. Appl. Remote Sens.*, vol. 11, no. 1, Jan. 2017, pp. 015005-1–15, doi: 10.1117/1.jrs.11.015005.

[17] 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, Jan. 2019.

[18] S. Mahajan and B. Fataniya, "Cloud detection methodologies: Variants and development—A review," *Complex Intell. Syst.*, vol. 6, no. 2, pp. 251–261, 2020.

[19] P. Li, L. Dong, H. Xiao, and M. Xu, "A cloud image detection method based on SVM vector machine," *Neurocomputing*, vol. 169, pp. 34–42, Dec. 2015, doi: 10.1016/j.neucom.2014.09.102.

[20] P. P. Joshi, R. H. Wynne, and V. A. Thomas, "Cloud detection algorithm using SVM with SWIR2 and tasseled cap applied to Landsat 8," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 82, Oct. 2019, Art. no. 101898, doi: 10.1016/j.jag.2019.101898.

[21] N. Ghasemian and M. Akhoondzadeh, "Introducing two random forest based methods for cloud detection in remote sensing images," *Adv. Space Res.*, vol. 62, pp. 288–303, Jul. 2018, doi: 10.1016/j.asr.2018.04.030.

[22] X. Chen, L. Liu, Y. Gao, X. Zhang, and S. Xie, "A novel classification extension-based cloud detection method for medium-resolution optical images," *Remote Sens.*, vol. 12, no. 15, 2020, Art. no. 2365, doi: 10.3390/RS12152365.

[23] L. Xu, R. Niu, S. Fang, and Y. Dong, "Cloud detection based on decision tree over Tibetan Plateau with MODIS data," in *Proc. 8th Int. Symp. Multispectral Image Process. Pattern Recognit. Int. Soc. Opt. Photon.*, 2013, vol. 8921, Art. no. 89210G, doi: 10.1117/12.2030399.

[24] P. Walder and I. Maclaren, "Neural network based methods for cloud classification on AVHRR images," *Int. J. Remote Sens.*, vol. 21, pp. 1693–1708, 2000, doi: 10.1080/014311600209977.

[25] M. J. Hughes and D. J. Hayes, "Automated detection of cloud and cloud shadow in single-date Landsat imagery using neural networks and spatial post-processing," *Remote Sens.*, vol. 6, no. 6, pp. 4907–4926, May 2014, doi: 10.3390/rs6064907.

[26] Q. Yuan, H. Shen, T. Li, Z. Li, and L. Zhang, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sens. Environ.*, vol. 241, Nov. 2020, Art. no. 111716.

[27] Z. Li, H. Shen, Q. Cheng, Y. Liu, S. You, and Z. He, "Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors," *ISPRS J. Photogrammetry Remote Sens.*, vol. 150, pp. 197–212, Apr. 2019, doi: 10.1016/j.isprsjprs.2019.02.017.

[28] X. Wu, Z. Shi, and Z. Zou, "A geographic information-driven method and a new large scale dataset for remote sensing cloud/snow detection," *ISPRS J. Photogrammetry Remote Sens.*, vol. 174, pp. 87–104, Apr. 2021.

[29] N. Ma, L. Sun, Q. Wang, Z. Yu, and S. Liu, "Improved cloud detection for Landsat 8 images using a combined neural network model," *Remote Sens. Lett.*, vol. 11, no. 3, pp. 274–282, Mar. 2020.

[30] L. Peng, L. Liu, X. Chen, J. Chen, X. Cao, and Y. Qiu, "Generalization ability of cloud detection network for satellite imagery based on DeepLabv3+," *Nat. Remote Sens. Bull.*, vol. 25, no. 5, pp. 1169–1186, 2021.

[31] M. Shi, F. Xie, Y. Zi, and J. Yin, "Cloud detection of remote sensingimages by deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2016, pp. 701–704.

[32] F. Xie, M. Shi, Z. Shi, J. Yin, and D. Zhao, "Multilevel cloud detection in remote sensing images based on deep learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3631–3640, Aug. 2017.

[33] G. Mateo-García, V. Laparra, L. P. Dan, and L. Gómez-Chova, "Transferring deep learning models for cloud detection between Landsat-8 and Proba-V," *ISPRS J. Photogrammetry Remote Sens.*, vol. 160, pp. 1–17, Feb. 2020.

[34] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3431–3440.

[35] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention*, 2015, pp. 234–241.

[36] 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, Aug. 2019.

[37] M. Wieland, Y. Li, and S. Martinis, "Multi-sensor cloud and cloud shadow segmentation with a convolutional neural network," *Remote Sens. Environ.*, vol. 230, pp. 111203, Sep. 2019, doi: 10.1016/j.rse.2019.05.022.

[38] J. Guo, J. Yang, H. Yue, X. Liu, and K. Li, "Unsupervised domain-invariant feature learning for cloud detection of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Oct. 2021, Art. no. 5405715, doi: 10.1109/TGRS.2021.3120001.

[39] J. Guo, J. Yang, H. Yue, and K. Li, "Unsupervised domain adaptation for cloud detection based on grouped features alignment and entropy minimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Mar. 2022, Art. no. 5603413, doi: 10.1109/TGRS.2021.3067513.

[40] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132–7141.

[41] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 3–19.

[42] J. Zhang, Y. Wang, H. Wang, J. Wu, and Y. Li, "CNN cloud detection algorithm based on channel and spatial attention and probabilistic upsampling for remote sensing image," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Aug. 2022, Art. no. 5404613, doi: 10.1109/TGRS.2021.3105424.

[43] 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, Feb. 2022, Art. no. 107890.

[44] "ATBD: Atmospheric correction algorithm: Spectral reflectances (MOD09) version 4.0," NASA contract NAS5-96062, 1999. [Online]. Available: https://modarch.gsfc.nasa.gov/data/atbd/atbd_mod08.pdf

[45] Z. Zhang, S. Yang, S. Liu, B. Xiao, and X. Cao, "Ground-based cloud detection using multiscale attention convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Aug. 2022, Art. no. 8019605, doi: 10.1109/LGRS.2021.3106337.

[46] E. F. Vermote and S. Y. Kotchenova, "MOD09 user's guide," 2008. [Online]. Available: http://modis-sr.ltdri.org

[47] R. C. Levy, S. Mattoo, L. A. Munchak, L. A. Remer, A. M. Sayer, and N. C. Hsu, "The collection 6 MODIS aerosol products over land and ocean," *Atmospheric Meas. Techn.*, vol. 6, no. 11, pp. 2989–3034, Nov. 2013, doi: 10.5194/amtd-6-159-2013.

Yawen He received the Ph.D. degree in geographic information science from the Institute of Geographic Sciences and Resources Research, Chinese Academy of Sciences, Beijing, China, in 2012.

His research interests include distributed sharing and integrated application of Marine environmental information.

Chuanxiang Dong is currently working toward the master's degree in mapping engineering with Shandong University of Science and Technology, Qingdao, China.

His research interests include computer vision and machine learning.

Nan Ma received M.Eng. degree in signal and information processing from Shandong Normal University, Jinan, China, in 2019. She is currently working toward the D.Eng. degree in energy and environmental protection from China University of Petroleum, Qingdao, China.

Her research interests include remote sensing and deep learning.

Yu Qu is currently working toward the master's degree in mapping engineering with Shandong University of Science and Technology, Qingdao, China.

His research interests include remote sensing image processing and deep learning.

Lin Sun received the Ph.D. degree in cartography and geographic information system from the Institute of Remote Sensing and Digital Earth, Chinese Academy of Science, Beijing, China, in 2006.

He is currently a Professor with the College of Geomatics, Shandong University of Science and Technology, Qingdao, China. His research interests include quantitative remote sensing and computer vision.

Huiyong Yu is currently working toward the Ph.D. degree in photogrammetry and remote sensing with Shandong University of Science and Technology, Qingdao, China.

His research interests include remote sensing image processing and deep learning.

Chenghu Zhou received the Ph.D. degree in cartography and GIS from the Institute of Geography, Chinese Academy of Sciences, Beijing, China, in 1992.

He is Cartographer and geographic information system scientist. He was elected academician of Chinese Academy of Sciences in 2013. He is currently a research fellow with the Institute of Geographic Sciences and Resources Research, Chinese Academy of Sciences. He has long been engaged in the research of remote sensing and GIS and its intersection with geographic sciences.