Research on Multi-temporal Cloud Removal Using D-S Evidence Theory and Cloud Segmentation Model

Xinwei Wang[1,*], Kailai Sun[1,*], Qianchuan Zhao[1,+], and Jianhong Zou[2,3]

[1] Center for Intelligent and Networked Systems, Department of Automation, BNRist,Tsinghua University,Beijing,100084,China
[2] Technology Development Department, Fujian Big Data Co., LTD, Fuzhou, China
[3] Fujian Nebula Big Data Application Service Co., LTD, Fuzhou, China

Abstract. Satellite remote sensing technology is widely applied for real-time drawing maps, forecasting weather, and predicting natural disasters. However, many remote sensing satellite images contain cloud noise, resulting in difficulties in practical applications. Existing cloud removal methods are limited when an image contains lots of clouds or thick clouds. Besides, most methods need a cloudless image as a reference. Our study proposes an advanced algorithm to remove cloud noise (especially thick clouds) in remote sensing images, including a cloud segmentation model, prior knowledge refinement, and Dempster-Shafer (D-S) evidence theory. Firstly, we trained a Cloud-net model to segment cloud regions. RGB and Near Infrared (NIR) images are fed into the Cloud-net model. Then it outputs coarse segmentation images. Secondly, we introduced the prior knowledge to refine and recover segmentation results. Finally, we designed cloud removal rules based on D-S evidence theory to fuse muti-temporal remote sensing images. Our method achieves a surprising performance on GaoFen-4 (GF-4) satellite images, reducing the average percentage of cloud noise from 30%-40% to 2%-8% per image.

Keywords: Satellite remote sensing image · D-S evidence theory · Cloud segmentation model · Multi-temporal denoising.

1 Introduction

Remote sensing is an essential tool for acquiring environmental information and natural resources. It perceives distant information about an object. Our study focuses on satellite remote sensing, which could be used for real-time mapping, weather forecasting, and natural disasters warning (e.g., typhoons and tsunamis). However, various kinds of noise will appear in remote sensing processes. How to denoise is a challenging and meaningful issue.

There are three main types of noise: impulsive noise caused by cosmic high-energy particle radiation; Gaussian noise generated by the transmission medium or the equipment; cloud noise caused by ground surface clouds and fog. Our study focuses on how to remove cloud noise. So far, there are more than 700

*These authors contributed equally to this work. +Corresponding author.
Emails: wxw21@mails.tsinghua.edu.cn; skl18@mails.tsinghua.edu.cn; zhaoqc@tsinghua.edu.cn; jianhongzou@foxmail.com.

remote sensing satellites all over the world that have captured a large number of images. However, lots of images cannot be directly used because of the heavy cloud noise. Removing the cloud noise can significantly increase the quality of existing remote sensing images.

Existing cloud removal studies can be divided into two categories: thin and thick cloud removal. These studies apply frequency domain filtering method, wavelet filtering method, etc., on a single image to remove thin clouds[1,2,3,4,5,6]. These methods rely on ground surface information but cannot work when thick clouds completely cover ground surface. Thick cloud removal studies can be divided into two categories: many methods are based on cloudless reference image [7,8,9]; and other methods fuse multi-temporal images [10,11,12]. The former methods select a cloudless image as the reference to reconstruct the target image. The target and the reference images are captured at the same location but at different times by the same satellite. The latter methods take multiple cloudy images, then repairs the target image by finding cloudless patches (or pixels) in reference images. These reference images are also captured at the same location but at different times, but they all contain clouds. These methods provide the possibility of removing thick clouds. However, the reference-based method needs cloudless images, which are hard to be obtained in practice. Besides, many of the mentioned methods are based on virtual synthetic cloudy images or few-cloud images, causing difficulties in practical applications.

To overcome these issues, we propose an accurate multi-temporal cloud removal method based on Cloud-net[13] and D-S evidence theory. Our approach is applied to natural satellite remote sensing image sequences. It performs well on natural heavy-cloudy images, not depending on cloudless images. It uses Cloud-net to conduct preliminary cloud segmentation on remote sensing images, then introduces color prior knowledge to refine and complement the segmentation maps, and finally applies the D-S evidence theory to fuse them.

To our best knowledge, previous cloud removal studies on natural images with a high percentage of cloud noise are rare. Our work has the following contributions: ① We trained Cloud-net and introduced color prior knowledge to improve cloud detection performance. ② Based on the D-S evidence theory, we designed a cloud removal rule that can effectively fuse multi-temporal remote sensing images. ③ We applied our method to real satellite remote sensing images and achieved a significant cloud removal performance. Compared with previous cloud removal methods, our method has the following advantages: ① we do not need cloudless images as the reference; ② we can apply our method to real remote sensing images containing thick clouds with excellent performance; ③ our method can deal with images with a high percentage of cloud noise.

2 Related work

2.1 Cloud removal methods

For the thin cloud removal methods, thin cloud regions are treated as a combination of ground surface and clouds. The thin clouds are removed by frequency domain filtering [1]. Authors [2] segment cloud regions using Poisson matting, then remove cloud noise using wavelet filtering. Other methods[3,4,5,6] treat

clouds as signals with specific characteristics and remove them through different filtering methods. These methods focus on processing a single thin cloudy image and cannot remove thick clouds.

The thick cloud removal methods [7,8,9,10,11,12] are developed from earlier research on blank pixel repairment of remote sensing images [14,15]. For the thick cloud removal methods, study [7] uses cloudless images of the same region as a reference, guiding the Markov random field (MRF) model to estimate the pixels in the target image. Studies [8,9] perform sparse decomposition of the cloudless image to obtain a dictionary with sparse coefficients to remove clouds. Authors [10] use a linear spectral decomposition method to recognize cloudy patches. Next, they use cloudless patches in different temporal images to fill the cloudy area. Study [11] is based on a global optimization process. When it detects a cloudy patch, it finds cloudless patches in different temporal images and then performs cloud removal using the two-dimensional gradient of cloudless regions. Study [12] performs cloud detection to extract cloudy patches. The cloudy and cloudless patches in different temporal images are fed into a deep convolutional network. At last, the network outputs cloud removal results. However, they have many limitations. Studies [7,8,9] use cloudless images as the reference. But in practice, it is hard to get them. Studies [7,8,9,12] use virtual synthetic cloudy images for training and testing, which is unreliable in practice. Study [10] uses an unreliable cloud detection method. The effect of cloud removal becomes poor when facing a large cloud coverage. Study [11] requires manual annotation of cloudy regions.

2.2 Cloud detection methods

Cloud detection methods are an essential part of our multi-temporal cloud removal method. They are mainly divided into three categories: ① Threshold methods [16]. ② Pattern recognition methods based on artificial neural networks [13,17,18,19,20]. ③ Spatial variation analysis methods and texture feature methods [21,22]. The threshold method sets a threshold to classify cloud and ground surface pixels by the difference in brightness and chromaticity. The spatial variation analysis and texture characterization methods are used to analyze the edge of the cloud to achieve more accurate and robust cloud detection. The primary neural network model for cloud detection is a semantic segmentation network based on U-net or its variants [23]. It can theoretically segment a single image's cloud targets at the pixel level. Other network structures also exist for cloud detection models, such as Cloud-net [13] and MSCN [19].

3 Methods

The aim of our study is to remove clouds on multi-temporal images, then reconstruct the true ground surface. Our method has four parts: ① Cloud-net for cloud segmentation. ② Color prior knowledge for refining segmentation results. ③ D-S evidence theory for combining multi-temporal segmentation results. ④ Evidence fusion method to reconstruct ground surface.

3.1 Cloud-net

We employee the convolutional neural segmentation network named Cloud-net [18], whose structure is shown in Fig. **??**. The network is divided into two

parts. The upper part extracts deeper features step by step. The lower part extracts multi-scale cloud information by concatenating the upper part features. It detects cloud regions accurately through its complex convolution module, not relying on pre-processing methods. The **contr_arm** module combines the original input with the output of the first and second convolution layers to produce a multi-scale feature map. Compared to the **contr_arm** module, **imprv_contr_arm** adds the third convolution layer; the **bridge** module adds a dropout layer; the **improve_ff_block** module concatenates the upper layer output with the lower layer output and obtains multi-scale feature information. At

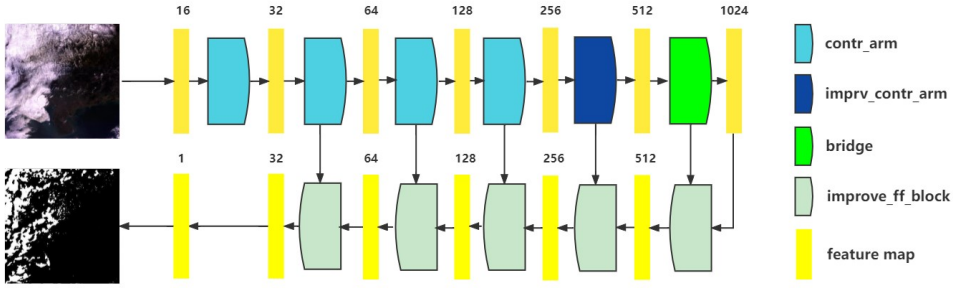

Fig. 1. Structure of the Cloud-net model

3.2 Color prior knowledge

There is a clear difference in satellite images covering large areas: images in cloud areas are whiter and brighter; images at ground surface/sea areas are bluer/darker. Based on this fact, to obtain better segmentation results of cloud areas, we convert the RGB image into the gray image[24] and then conduct Eq. 1 on the cloud confidence map generated by Cloud-net.

$$Conf'(i,j) = Conf(i,j) + J[Gray(i,j), L] \times Bias,$$
$$J[Gray(i,j), L] = \begin{cases} 1, & Gray(i,j) > L \\ 0, & Gray(i,j) \leq L, \end{cases} \tag{1}$$

where *Conf(i,j)* is the cloud confidence value at location *(i,j)* output by Cloud-net model. $Conf'(i,j)$ is the refined confidence value. *Gray(i,j)* is the gray value at location *(i,j)*. J is a function to classify pixels in the gray image. L is a fixed classification threshold in J. *Bias* is a fixed correction value for pixels that are classified bright by J. Here, $Conf'(i,j)$ may exceed 1. We will talk about how to handle this in Section 3.3

This prior knowledge effectively classifies cloud regions on large-scale remote sensing images. For small-scale objects (e.g., white buildings) that may bring errors, our method increases same confidence for these objects and clouds. In this situation, their classification results only rely on cloud-net. Thus in the worst case, the performance with the method is no worse than without it.

3.3 D-S evidence theory

The D-S evidence theory originated in the 1960s and was proposed by A.P. Dempster and his students [25]. The theory is based on uncertain evidence and performs information fusion to produce reliable results.

We apply the theory of migration to the cloud removal problem with the following definitions: at the time t, the cloud-net outputs cloud confidence 0.86 in the location *(i,j)*, then:

$$m_{(i,j)t}(C) = 0.86, \tag{2}$$

where m means the mass function. C means event 'cloud'. The Cloud-net outputs cloud confidence for each location, which is the mass function value. Due to color prior knowledge, cloud confidences may exceed 1. We then normalize them through dividing them by the biggest one. Considering that Cloud-net is not entirely reliable, we add an uncertainty value U to represent its unreliability. At this point, our identification framework has three judgments for a given location: cloud, cloudless, and uncertain (U). The sum of cloud confidence and cloudless confidence in the final calculation is 1 - U, as Eq. 3.

$$\begin{aligned} m_{(i,j)t}(C) + m_{(i,j)t}(CL) &= 1, \\ m'_{(i,j)t}(C) + m'_{(i,j)t}(CL) &= 1 - U, \end{aligned} \tag{3}$$

where CL represent event 'cloudless'. m' means mass function value after introducing U. Because there are only two assertions: the overall multi-temporal information is cloudy (OC) or cloudless (OCL). Our study only used the plausibility function in D-S evidence theory to calculate the final confidence based on multi-temporal confidence information as in Eq. 4.

$$\begin{aligned} Pla\,(OC) &= \frac{1}{K} \sum_{A_1 \cap A_2 \ldots \cap A_q = C} m_1\,(A_1)\, m_2\,(A_2) \ldots m_q\,(A_q)\,, \\ Pla\,(OCL) &= \frac{1}{K} \sum_{A_1 \cap A_2 \ldots \cap A_q = CL} m_1\,(A_1)\, m_2\,(A_2) \ldots m_q\,(A_q)\,, \end{aligned} \tag{4}$$

where *Pla* represents the value of the plausibility function supporting a particular assertion. K is the normalization factor. A_t represents possible events (cloudy, cloudless, and uncertain) at time t. $m_t(A)$ represents mass function value of event A at time t. The representation of location *(i,j)* is omitted for brevity.

Here, we specifically define that:

$$C \cap U = C, \quad CL \cap U = CL, \quad C \cap CL = \emptyset. \tag{5}$$

To get an assertion, we use this rule to perform evidence fusion on multi-temporal cloud confidence maps. Then we use the assertion and multi-temporal images to implement cloud removal.

3.4 Evidence fusion method

After obtaining the refined confidence map and fusing the results using D-S evidence theory, we fuse the data based on the following rules.

If the D-S evidence theory fusion result in the location *(i,j)* is OCL, then:

Step1. Takes the RGB values of pixels $P_1, P_2 ... P_m$ in time series 1,2,...m at location *(i,j)* for clustering.

Step2. If there exists a pixel clustering center C of which grayscale value is less than the brightness threshold L, and the number of pixels belonging to C exceeds threshold N_1, we select C as the fusion result. If there is no such clustering center, go to Step3.

Step3. Select the top N_2 pixels with the lowest cloud confidence, then take their mean value as the fusion result.

If the fusion result is OC, reduce N_1 and N_2. Then conduct the above steps.

N_1 and N_2 are defined based on the length of time series. We design this rule based on the assumption that these multi-temporal images have slight ground surface variation. Cloud differs sea/ground significantly. Therefore if there are multiple dark pixels with similar RGB values at the same location, these pixels could represent the true ground surface better. If there are no such pixels, we can only trust the Cloud-net and the prior knowledge, then take the mean value of the top N_2 pixels with the lowest cloud confidence as output.

4 Experiment

First, we trained Cloud-net in Section 4.1 to figure out whether it can perform well on OOD (Out of Distribution) dataset. Second, we designed an experiment to prove that color prior knowledge can refine segmentation results of the Cloud-net in Section 4.2. Third, we show that our method can effectively remove the cloud and reconstruct the ground surface in Section 4.3. Then. we designed ablation experiment to discuss the effects of the prior knowledge and the D-S evidence theory in Section 4.4. Finally, we demonstrate that our method is superior to other existing cloud-removal methods in Section 4.5. The experiment was run in the following environment: software environment: Python 3.6, Tensorflow 1.12.0, Keras 2.2.4, Scikit-image 0.15.0, windows 10; hardware environment: RTX3090, AMD EPYC 7702 processor.

4.1 Training Cloud-net

We selected the dataset and pre-trained network weights provided by [13]. The dataset contains 8400 images (resolution: 384×384) from the Landsat 8 satellite. After removing the images whose blank areas are more than 10%, it remains 4382 images with four channels (RGB and NIR). We use a transfer learning strategy to fine-tune the pre-trained weights. In our study, we choose 10^{-5} as the learning rate, 20% of the total dataset as the validation set, 80% as the training set, 16 as batchsize, and other parameters are set to default values. The loss function is defined as shown in Eq. 6.

$$loss = 1 - \frac{sum(\boldsymbol{y_{pred}} \odot \boldsymbol{y_{gt}}) + \varepsilon}{sum(\boldsymbol{y_{pred}} + \boldsymbol{y_{gt}}) - sum(\boldsymbol{y_{pred}} \odot \boldsymbol{y_{gt}}) + \varepsilon}, \tag{6}$$

where $\boldsymbol{y_{pred}}$ is the predicted 384×384 matrix, $\boldsymbol{y_{gt}}$ is the labeled matrix. $\odot$ is the Hadamard product. *sum()* is matrix summation function. ε is a small value that is larger than 0 (we choose 10^{-7} in the experiment). The loss function value approaches convergence quickly. The test images are part of the data taken by GF-4. There are 34 images spread over six locations, each containing 4 to 7 images at different times. Part of the test images and corresponding cloud segmentation results are shown in Fig.2. The network has an adequate cloud detection performance but demonstrates poor performance when facing OOD data from the training set shown in Fig. 2. So we use the color prior knowledge to refine and complement segmentation results.

4.2 Experiments with color prior knowledge

To make the segmentation results more accurate, we introduce the color prior knowledge module in Section 3.2. Experimental results are shown in Fig. 2. The color prior knowledge module effectively improves the segmentation results when facing OOD data. Using manually annotated cloud pixels as ground truth, the color prior knowledge can increase the segmentation accuracy defined from an average of 30% to over 90%. Segmentation accuracy (SA) is calculated by $SA = AN/TN$, where AN is the number of accurately classified cloud pixels, TN is the number of total cloud pixels.

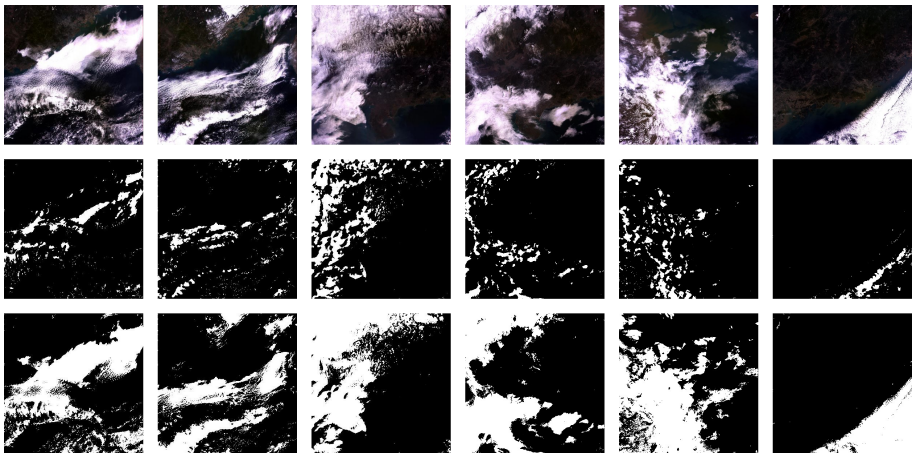

Fig. 2. Comparison of input against the output of the Cloud-net with the color prior knowledge. Row 1: test images captured by GF-4 satellite. Row 2: segmentation results of Cloud-net, and the white part are cloud regions classified by cloud-net. Row 3: refined segmentation results by color prior knowledge.

4.3 Multi-temporal remote sensing cloud removal experiments

To verify the effectiveness of our cloud removal method, we took multi-temporal images captured by the GF-4 satellite as test data. Part of the experimental results is shown in Fig. 3. As our method tends to select darker pixels and the image itself is not very bright, we conducted a gamma transform on reconstructed images to make them clearer. As we can see, the final output image of the algorithm contains very few clouds. To measure its performance quantitatively, we manually annotate the cloud areas in test images. We calculate the cloud rate in each input image and corresponding reconstructed results. The cloud rate (CR) is calculated by this: $CR = C_n/T_n$, where C_n is the number of cloud pixels, T_n is the number of total pixels. As shown in Table. 1 and Fig. 3, our method removes most clouds. However, few cloud areas still exist after cloud removal. The main reason is that the areas are always covered with clouds in an image sequence. Therefore, there is no true ground surface information available for reconstruction. On the other hand, the above results show that more images captured at the same location and more complementary cloudless areas will lead to better cloud removal results.

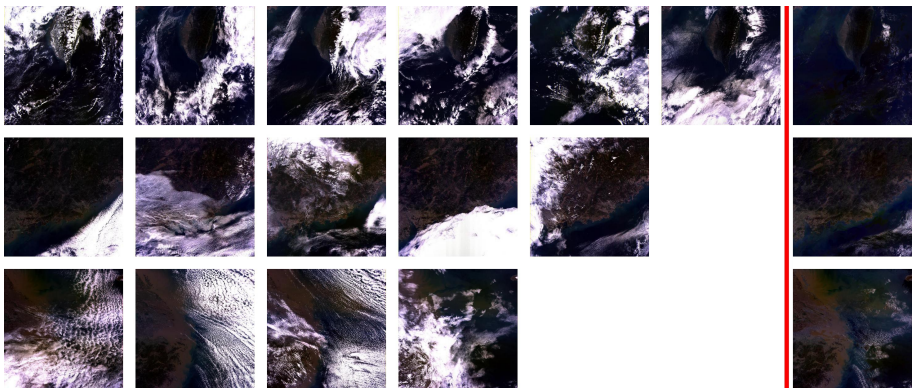

Fig. 3. Results of our cloud removal method. The last column is the final cloud removal results, and the other columns are the multi-temporal remote sensing images.

Table 1. Estimated cloud rates before and after cloud removal. Minimum and Maximum means the smallest and biggest cloud rates in these areas. WLR, STS-CNN and PSTCR will be introduced in Section 4.5

Area/Cloud rate	Minimum	Maximum	WLR	STS-CNN	PSTCR	Our method
1	26.65%	41.96%	22.51%	23.07%	20.49%	**2.55**%
2	16.30%	44.31%	33.90%	15.44%	24.59%	**2.66**%
3	30.12%	49.02%	28.35%	26.89%	22.76%	**2.14**%
4	40.65%	46.23%	33.86%	24.74%	29.79%	**7.37**%
5	21.31%	48.68%	31.13%	26.75%	25.85%	**2.67**%
6	24.34%	50.31%	31.79%	29.32%	17.51%	**1.76**%

4.4 Ablation experiment

We designed an ablation experiment to verify the effect of the prior knowledge and the D-S evidence theory. To remove the D-S evidence theory, we simply output the pixel with the lowest cloud confidence. We did not remove Cloud-net because it is essential. The results of ablation experiment are shown in Fig. 4 and Table. 2.

Table 2. Cloud rates in ablation experiment. *W/O_P* means our method without the prior knowledge. *W/O_D-S* means without our method the D-S evidence theory. *ALL* means our method with all components.

Area/Ablation	*W/O_P*	*W/O_D-S*	*ALL*
1	28.46%	19.81%	**2.55**%
2	15.69%	33.20%	**2.66**%
3	24.31%	31.38%	**2.14**%
4	39.65%	20.80%	**7.37**%
5	33.11%	19.27%	**2.67**%
6	19.27%	14.73%	**1.76**%

We can see from Fig. 4 and Table. 2: outputs of *W/O_P* still contain lots of clouds, which means the prior knowledge can significantly refine the results of cloud segmentation. Outputs of *W/O_D-S* contain less cloud than *W/O_P* overall, but more cloud than *ALL*. Also, we found that the outputs of *W/O_D-S* contain many cloud pixels which are similar to salt noise. Obviously, the D-S evidence theory can make outputs more smooth.

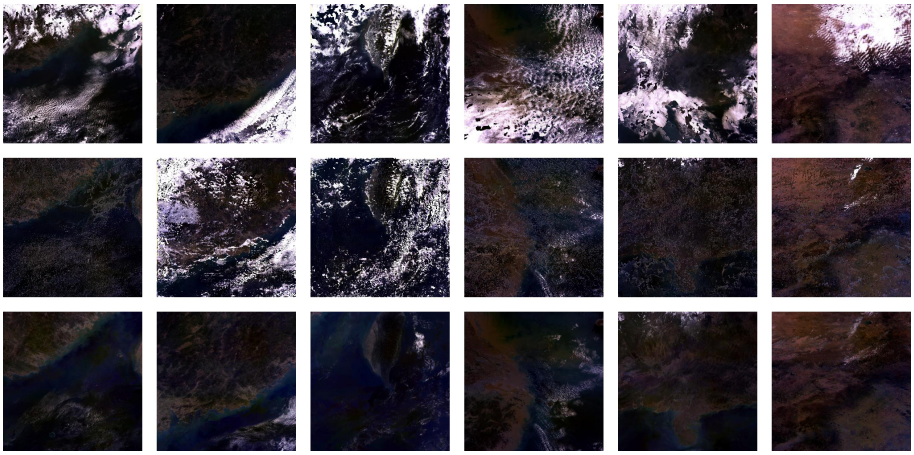

Fig. 4. Row 1: outputs without the prior knowledge. Row 2: outputs without the D-S evidence theory. Row 3: outputs with all components.

4.5 Comparison against other cloud removal methods

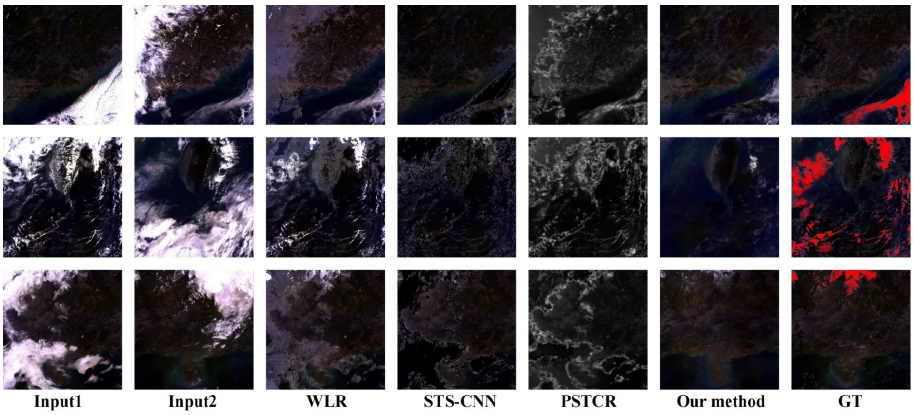

Fig. 5. Column 1 and column 2: inputs. Column 3: outputs of WLR. Column 4: outputs of STS-CNN. Column 5: outputs of PSTCR, which can only process grayscale image. Column 6: outputs of our method. Column7: ground truth (combination of cloudless parts from two inputs).

For previous thick cloud removal methods, we choose WLR [14], STS-CNN [15], and PSTCR [12] for comparison. WLR and STS-CNN take cloudy target images, cloudless reference images, and cloud segmentation maps as input. Then they use the reference images to fill the target images. PSTCR can only process gray-scale images, selects multiple cloudy images with cloud segmentation maps as input, and supplements the images with a convolutional neural network. After obtaining the codes corresponding to the above methods, we followed its original hyper-parameters and experimented with manually labeled cloudy images as inputs. To quantify the cloud-removal effect, we fused their inputs by combining their cloudless regions to obtain the complementary image *gt*. The red regions in *gt* mean both regions in two inputs are cloudy. Subsequently, the *MSE*(Mean

Table 3. Comparison of MSE against different cloud removal methods on test images

Area/Methods	WLR	STS-CNN	PSTCR	Our method
1	0.0552	0.0931	0.0844	**0.0501**
2	0.1152	0.0200	0.1204	**0.0191**
3	1.066	0.8268	0.9092	**0.7813**
4	0.1484	0.0886	0.2025	**0.0335**
5	1.115	0.7198	0.8202	**0.6664**
6	0.0832	0.1190	0.0416	**0.0348**

Square Error) was calculated for the cloudless part of *gt* and the corresponding part of the output images processed by different methods, defined as in Eq. 7.

$$MSE = \frac{\sum_{(i,j)\in CFR} (G_{(i,j)} - R_{(i,j)})^2}{T}, \tag{7}$$

where *(i,j)* $\in CFR$ represents the coordinates of the cloudless part, G represents the complementary image after pixel normalization, R represents the output of each method after cloud removal, and T represents the total number of pixels in the cloudless region.

As can be seen from Table. 1 ,Table. 3 and Fig. 5, our method outperforms other methods. These methods use manual annotations, while our method does not need manual annotation. In the case of realistic multi-temporal massive cloudy images, the *MSE* values output by our method are lower than those of the other methods. Furthermore, because we use the original image's pixels (or clustering centers) directly for filling, our restored images are smoother and more realistic without producing many blank regions and mutations, as shown in Fig. 5.

5 Conclusion

Our study presents a multi-temporal thick cloud removal method, applying cloud segmentation networks, prior knowledge, and D-S evidence theory. Our method does not require cloudless images as the reference. Reconstructed images are close to the true ground surface. Our method significantly reduces the cloud rate on GF-4 satellite remote sensing images. Theoretically, our method can achieve complete cloud removal when thick clouds cover not all the regions at the same location. However, our study still has limitations: Our study does not consider dealing with the shadows from cloud occlusion now. Thin clouds are not easily recognized in our study. Our method is based on the assumption that the true ground surfaces in different images differ very little. If the overall brightness or true ground surface changes significantly, they may lead to wrong fusion results. These problems need to be addressed in future research.

Acknowledgement: This work is supported by Key RD Project of Chinaunder Grant No.2017YFC0704100, 2016YFB0901900, National Natural Science Foundation of China under Grant No. 61425024, the 111 International Collaboration Program of China under Grant No.BP2018006, 2019 Major Science and Technology Program for the Strategic Emerging Industries of Fuzhou under Grant No.2019Z-1, and in part by the BNRist Program under Grant No.BNR2019TD01009.

References

1. Liu, J., Wang, X., Chen, M., Liu, S., Zhou, X., Shao, Z., Liu, P.: Thin cloud removal from single satellite images. Optics express **22**(1), 618–632 (2014) 1, 2.1
2. Wang, Z., Jin, J., Liang, J., Yan, K., Peng, Q.: A new cloud removal algorithm for multi-spectral images. In: MIPPR 2005: SAR and Multispectral Image Processing. vol. 6043, p. 60430W. International Society for Optics and Photonics (2005) 1, 2.1
3. Xu, M., Pickering, M., Plaza, A.J., Jia, X.: Thin cloud removal based on signal transmission principles and spectral mixture analysis. IEEE Transactions on Geoscience and Remote Sensing **54**(3), 1659–1669 (2015) 1, 2.1
4. Li, J., Wu, Z., Hu, Z., Zhang, J., Li, M., Mo, L., Molinier, M.: Thin cloud removal in optical remote sensing images based on generative adversarial networks and physical model of cloud distortion. ISPRS Journal of Photogrammetry and Remote Sensing **166**, 373–389 (2020) 1, 2.1
5. Chun, F., Jian-wen, M., Qin, D., Xue, C.: An improved method for cloud removal in aster data change detection. In: IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium. vol. 5, pp. 3387–3389. IEEE (2004) 1, 2.1
6. Tarel, J.P., Hautiere, N.: Fast visibility restoration from a single color or gray level image. In: 2009 IEEE 12th International Conference on Computer Vision. pp. 2201–2208. IEEE (2009) 1, 2.1
7. Cheng, Q., Shen, H., Zhang, L., Yuan, Q., Zeng, C.: Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal mrf model. ISPRS journal of photogrammetry and remote sensing **92**, 54–68 (2014) 1, 2.1
8. Xu, M., Jia, X., Pickering, M., Plaza, A.J.: Cloud removal based on sparse representation via multitemporal dictionary learning. IEEE Transactions on Geoscience and Remote Sensing **54**(5), 2998–3006 (2016) 1, 2.1
9. Li, X., Wang, L., Cheng, Q., Wu, P., Gan, W., Fang, L.: Cloud removal in remote sensing images using nonnegative matrix factorization and error correction. ISPRS journal of photogrammetry and remote sensing **148**, 103–113 (2019) 1, 2.1
10. Tseng, D.C., Tseng, H.T., Chien, C.L.: Automatic cloud removal from multi-temporal spot images. Applied Mathematics and Computation **205**(2), 584–600 (2008) 1, 2.1
11. Lin, C.H., Tsai, P.H., Lai, K.H., Chen, J.Y.: Cloud removal from multitemporal satellite images using information cloning. IEEE transactions on geoscience and remote sensing **51**(1), 232–241 (2012) 1, 2.1
12. Zhang, Q., Yuan, Q., Li, J., Li, Z., Shen, H., Zhang, L.: Thick cloud and cloud shadow removal in multitemporal imagery using progressively spatio-temporal patch group deep learning. ISPRS Journal of Photogrammetry and Remote Sensing **162**, 148–160 (2020) 1, 2.1, 4.5
13. Mohajerani, S., Saeedi, P.: Cloud-net: An end-to-end cloud detection algorithm for landsat 8 imagery. In: IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium. pp. 1029–1032. IEEE (2019) 1, 2.2, 4.1
14. Zeng, C., Shen, H., Zhang, L.: Recovering missing pixels for landsat etm+ slc-off imagery using multi-temporal regression analysis and a regularization method. Remote Sensing of Environment **131**, 182–194 (2013) 2.1, 4.5
15. Zhang, Q., Yuan, Q., Zeng, C., Li, X., Wei, Y.: Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network. IEEE Transactions on Geoscience and Remote Sensing **56**(8), 4274–4288 (2018) 2.1, 4.5

16. Derrien, M., Farki, B., Harang, L., LeGleau, H., Noyalet, A., Pochic, D., Sairouni, A.: Automatic cloud detection applied to noaa-11/avhrr imagery. Remote Sensing of Environment **46**(3), 246–267 (1993) 2.2
17. Jang, J.d., Viau, A.A., Anctil, F., Bartholomé, E.: Neural network application for cloud detection in spot vegetation images. International Journal of Remote Sensing **27**(4), 719–736 (2006) 2.2
18. Mohajerani, S., Krammer, T.A., Saeedi, P.: Cloud detection algorithm for remote sensing images using fully convolutional neural networks. arXiv preprint arXiv:1810.05782 (2018) 2.2, 3.1
19. Li, Z., Shen, H., Wei, Y., Cheng, Q., Yuan, Q.: Cloud detection by fusing multi-scale convolutional features. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci **4**, 149–152 (2018) 2.2
20. Ma, N., Sun, L., Wang, Q., Yu, Z., Liu, S.: Improved cloud detection for landsat 8 images using a combined neural network model. Remote Sensing Letters **11**(3), 274–282 (2020) 2.2
21. Solvsteen, C.: Correlation-based cloud detection and an examination of the split-window method. In: Global Process Monitoring and Remote Sensing of the Ocean and Sea Ice. vol. 2586, pp. 86–97. SPIE (1995) 2.2
22. Chen Peng, Zhang Rong, L.Z.: Feature extraction in remote sensing image cloud map recognition. Journal of the University of Science and Technology of China (5), 234–241 (2009) 2.2
23. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015) 2.2
24. Kumar, T., Verma, K.: A theory based on conversion of rgb image to gray image. International Journal of Computer Applications **7**(2), 7–10 (2010) 3.2
25. Yager, R.R.: On the dempster-shafer framework and new combination rules. Information sciences **41**(2), 93–137 (1987) 3.3