

# A Novel Dense-Attention Network for Thick Cloud Removal by Reconstructing Semantic Information

Yuyun Chen , Zhanchuan Cai , *Senior Member, IEEE*, Jieyu Yuan , and Lianghai Wu 

**Abstract**—The presence of thick clouds in single optical images shows the contamination of interesting objects. Besides, the difficulty of thick cloud removal is mainly the restoration of the weak boundary information from cloud-contaminated areas. Recently, many deep-learning-based frameworks were applied for cloud removal by obtain the related semantic information from the weak boundary information. However, the large-size cloud-contaminated areas lead to the artificial textures in the resulting images. Thus, obtaining the optimal semantic information from finite boundary information is the key to solve this problem. In this work, we design a deep-learning framework for cloud removal, especially large-size clouds removal (i.e., more than 30% coverage of the whole image). First, we design a cloud location model (CLM), which adopted the fully convolutional network to locate the cloud. Second, desired by theory of the coarse-to-fine restoration, we build a dense-attention network (termed as DANet) for restoring cloud contaminated areas. In the DANet, we design a dense block into the coarse network for training the features of restoring directions of each pixel from the weak boundary information. Furthermore, a contextual attention module is built into refinement network for restoring contaminated areas relying on the semantic relationship between the background and foreground information. Compared with the state-of-the-art methods, the proposed DANet achieved greater removal performances and reconstruct more natural image textures.

**Index Terms**—CLM, contextual attention module, DANet, dense block, thick cloud removal.

## I. INTRODUCTION

REMOTE sensing images obtained by space-borne satellites usually contain clouds, causing the ground information to be easily disturbed by clouds [1]. According to statistics, 35% of the Earth's surface is covered by clouds during the year [2]. The cloud images affect our reading of the ground truth and limit their application. Therefore, we need to remove these clouds from the images, which are contaminated with cloud.

Manuscript received 29 October 2022; revised 23 December 2022; accepted 9 January 2023. Date of publication 11 January 2023; date of current version 2 March 2023. This work was supported in part by the Science and Technology Development Fund of Macau under Grant 0025/2019/AKP, Grant 0052/2020/AFJ, and Grant 0059/2020/A2, in part by the National Foreign Experts Project—Senior Foreign Experts Introduction Program—Digital data processing method and key theory of smart factory based on digital twin, and in part by Zhuhai Industry-University-Research Collaboration Program under Grant ZH22017002210011PWC. (Corresponding author: Zhanchuan Cai.)

Yuyun Chen, Zhanchuan Cai, and Jieyu Yuan are with the School of Computer Science and Engineering, Macau University of Science and Technology, Macau 999078, China (e-mail: yuyunchen\_cn@foxmail.com; zccai@must.edu.mo; jieyuyuan.cn@gmail.com).

Lianghai Wu is with the School of Computer Science and Engineering, Macau University of Science and Technology, Macau 999078, China, and also with the School of Computer Science, Guangdong University of Petrochemical Technology, Maoming 525000, China (e-mail: wlh@gdupt.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2023.3236384

In remote sensing image cloud removal, it is general divided into thin and thick cloud removal [3]. Compared with thin clouds, thick clouds have stronger occlusion and less ground information, and it is challenging to restore the information of the original ground image. Therefore, thick cloud removal has received considerable attention in recent cloud removal studies [4].

At present, thick cloud removal methods are mainly divided into auxiliary-data-based methods and individual-based methods [5]. The auxiliary-data-based methods rely on clouds state information obtained at different time points or multiband information obtained by different sensors to recover the ground truth information [6]. To referring more details of contaminated areas, many algorithms restored the ground truth information combined with some assisted information (e.g., time-sequenced image information [4], [7], image textural information from multitemporal sensors [8], [9], [10]). However, in some case (e.g., harsh climates), single images potentially were the only source data [14]. Thus, cloud removal from single images is still changing and important. Therefore, thick cloud removal from a single image is still variable and important.

Individual-based cloud removal relies on spatial information for inpainting, which relies on limited boundary information of occluded objects on thick cloud boundaries. In traditional methods, thick cloud removal are mainly based on interpolation, filling of surrounding sample blocks [11], [12]. The work of [13] and [14] is mainly based on different methods to remove cloud, but they are only suitable for removing remote sensing images with regular surrounding textures. Lorenzi et al. [15] proposed a sample patch-based cloud removal method, which mainly reconstructs missing areas in a given image by propagating spectral geometric information extracted from the rest of the image. However, these methods can only repair the simple texture blocks. The semantic information recovered is far from the original.

Recently, deep learning-based methods have become popular in cloud removal studies and achieved better performance. Zheng et al. [16] proposed an adversarial generative network (GAN) framework for cloud removal, but the boundaries are easily blurred and unclear after cloud removal. Zhang et al. proposed spatial-temporal-spectral deep convolutional neural network (STS-CNN) [17] to repair different types of missing information, including deadlines in Aqua MODIS band 6, Landsat SLC-off problem, and thick cloud removal problem. Besides, Zhang et al. proposed a network based on a combination of deep spatiotemporal prior and low-rank tensor singular value

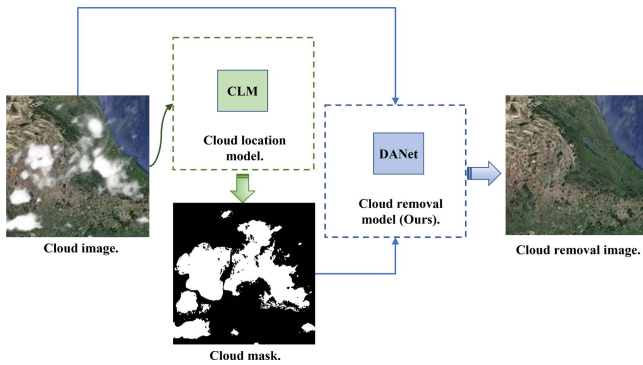


Fig. 1. Cloud removal framework. The proposed DANet has two main components: CLM and cloud removal model. We segment the cloud with CLM to local the cloud in the image. The output of CLM as location mask to guide the DANet for coarse-to-fine cloud removal.

decomposition (DP-LRTSVD) [18], which can effectively remove thick clouds by combining model-driven and data-driven strategies. Xu et al. proposed an attention mechanism-based generative adversarial networks for cloud removal in Landsat images (AMGAN-CR) [19] by using an attention map guided attention residual network, which can effectively recover the semantic information of the ground.

However, the current individual-based methods to remove the cloud still have the following problems:

1) Traditional methods use only semantic information around to make prediction of ground information, which can achieve better results for simple and regular ground information, but not ideal for removing cloud contaminated areas with complex textures.

2) Although the method based on deep learning achieves better results in cloud removal, the effect of remote sensing image removal is not particularly stable for remote sensing images contaminated by large areas of clouds (covering more than 30% of the whole image area). The restored semantics has a certain gap with the original semantic information on the ground and is prone to certain artifacts.

Therefore, a method applicable to large-scale cloud removal is the key to fill the research gap of cloud removal.

To recover the large area occluded by clouds (more than 30%), we propose a complete cloud removal framework: cloud location and cloud removal, as shown in Fig. 1. First, the thick cloud is located by the cloud location model (CLM) by adopting the FCN. Second, in the cloud removal part, the semantic deviation or artifacts are easily produced in large-size cloud removal. The long-distance correlation between regions is ignored when the ordinary deep network restores the cloud. Therefore, we proposed DANet based on a coarse-to-fine restoration framework to remove the cloud. In the coarse stage, we design a dense block to strengthen the propagation of semantic features. In the fine stage, we introduce a contextual attention module to improve the semantic connection between the missing regions and the overall image. DANet removes the cloud from coarse to fine in all aspects, so as to restore the original semantic information of the ground as much as possible.

So far, synthetic cloudy images created based on the atmosphere scattering model are one of the common test methods for evaluating the quantify of the deep-learning frameworks for cloud removal [6]. In this work, we use NMPU-RESISC45 [20] and NOAA [21] datasets as training and testing dataset. The total number of the training dataset is 35 000. A total of 24 500 images of the dataset are applied for training, and the rest of them are used to create synthetic cloudy images as the testing dataset by using the common atmosphere cloud model [22], [23].

The main contributions of this article are summarized as follows:

1) *A framework for cloud removal is proposed.* We propose a two-stage framework for cloud removal: CLM and cloud removal model. Among them, we propose the CLM to locate and extract the cloud areas more accurately, which facilitate the accuracy of thick cloud removal.

2) *The thick cloud removal network is proposed.* Based on the coarse-to-fine restoration, the DANet is proposed to remove thick cloud. It solves the problems of insufficient reduction degree and artificial artifacts in the restoration using only limited boundary information of the masked object, and comprehensively and finely restores the objects masked by thick clouds.

3) *Achieving the better recovery effect of large-size thick cloud.* In the DANet, a residual module and a contextual attention module are designed based on a coarse-to-fine network, which improve the association between the cloud-occluded regions and the existing semantics. Therefore, the ground information occluded by clouds can be recovered more accurately by using more limited boundary information.

The rest of this article is organized as follows. The details of the cloud removal framework is introduced in Section II. Section III presents the experiments. Section IV gives the discussion of the results. The conclusion of this work is stated in Section V.

## II. METHOD

For the thick cloud removal with more than 30% coverage of the whole image, we propose a two-stage thick cloud removal framework to make the cloud removed area closer to the context of the cloudless state using the limited boundary information. In the first stage, the CLM is proposed to locate extra thick cloud. In the second stage, the DANet is proposed to remove the thick cloud areas. The specific details are as follows:

### A. Cloud Removal Strategy

A complete cloud removal process includes cloud location and cloud removal network.

In the first stage of cloud removal, the cloud and background image in the multispectral remote sensing image is extracted to the locate the cloud. In the CLM, we adopt the fully convolutional networks (FCN) [24] to segment and extract thick clouds, which locate the cloud area. In the FCN, the simple linear iterative clustering (SLIC) [25] is used to perform superpixel preprocessing on the image. Superpixel is a small region consisting of a series of adjacent pixels with similar characteristics such as color, brightness, texture, etc. Therefore, FCN expresses the

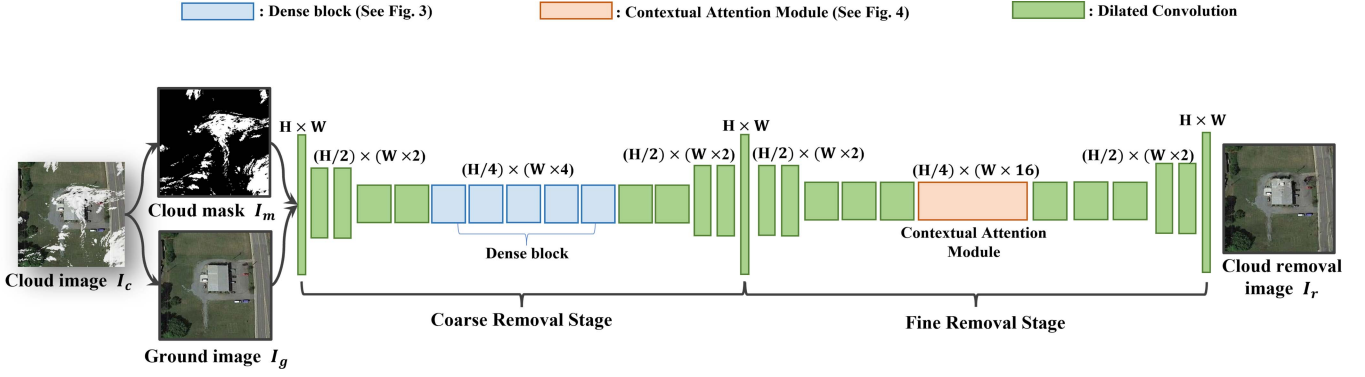


Fig. 2. Framework of the DANet that is divided into coarse and refinement network. In the coarse cloud removal part, a dense block is designed. In the refinement network, we adopt the contextual attention module.

characteristics of the superpixel area by selecting the key pixels. The superpixel areas must be distinguished as cloud or noncloud by identifying the characteristics of the key pixels through the network. After the real cloud image is segmented by FCN, the areas of the cloud are located, and the cloud mask image is generated.

In the second stage of cloud removal framework, to better recover the information of the area occluded by the cloud, we designed the DANet. For DANet, it needs to input cloud image and the cloud accurate location (cloud mask), and finally output the recover cloud-free image. The process is shown in the blue part of Fig. 1.

In the cloud removal framework, we first locate the extra cloud image  $I_c$  into a ground image  $I_g$  and cloud mask  $I_m$  by CLM. Then,  $I_m$  and  $I_g$  are trained as input of the DANet, and the cloud removal image  $I_r$  is the output. We define the DANet as the  $G_m$ , the output cloud removal image  $I_r$  is as follows:

$$I_r = I_g \odot I_m + G(I_c, I_m) \odot (1 - I_m) \quad (1)$$

where  $\odot$  is the elementwise product. In all experiments, the image values are linearly scaled to  $[-1, 1]$ . On the training dataset, the network  $G_m$  is updated and repaired using the spatial discount loss on each epoch.

## B. DANet

The existing cloud removal network can not recover the semantics of a single remote sensing image well, especially for cloud coverage over than 30%. To improve the semantic information of the original image, we propose DANet for thick cloud removal based on coarse-to-fine restoration. In the network structure of DANet, a dense block is designed in the coarse network, and contextual attention module is combined in the fine network. The flowchart of the DANet is shown in Fig. 2.

1) *Coarse-to-Fine Network*: We mainly remove the thick cloud coverage more than 30% for the single remote sensing images, the recovered area is relatively large. The common cloud removal network usually relies on the semantic information of surrounding regions, while ignoring the long-distance correlation between regions, which is prone to semantic deviation or artifacts after cloud removal [26]. Therefore, our proposed

network (DANet) is based on a coarse-to-fine restoration framework. With two-stage progressive generation, the framework will make the recovered images more realistic and perceptually consistent.

In the coarse removal stage, to enhance the propagation of semantic features, we add a dense block, which realize the approximate information of the area occluded by the cloud can be trained from the existing semantics. In the fine removal stage, to overcome the situation that convolutional networks cannot effectively learn from distant spatial image features, a context attention module is introduced to improve the semantic connection between the missing region and the overall image in the deep generative network.

In terms of layer implementation, we use mirror padding for all convolutional layers and remove the batch normalization (BN) layer [27]. The BN layer could cause the network to be unstable due to the inappropriate parameter settings of different networks [28]. Therefore, we remove the BN layer in this network to improve the robustness of the network. In addition, we use exponential linear element (ELU) [29] as the activation function and use clipping operation to output the filtered values. Therefore, the mean of the output in the ELU activation is close to zero.

2) *Dense Block*: In the first step of coarse recover, we need to control the semantic direction to facilitate the next step of fine repair. Since the cloud coverage is more than 30% that relatively large, the distance between the cloud region and the other region is far, it is necessary to deepen the network for training to improve the recovery of semantic information.

However, in the network, the information of each layer can only be shared with the information of the previous layer. As the depth of the training network increases, the information with larger layers is more disconnected from the earlier information. As long as there is a deviation in the training of one layer, it is easy to accumulate later, resulting in a greater deviation. Thus, the semantics are prone to distortion and artifacts. To solve the problem of communication between information in deep networks, we design dense blocks in the network.

Dense block enables all layers to easily access previously computed feature maps and effectively alleviates the vanishing gradient problem [30]. The dense connection improves feature

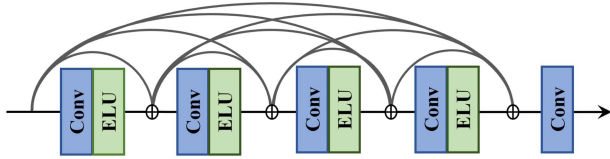


Fig. 3. Dense block. Each dense block consists of five parts, each built with a stacked convolution and ELU activation function, except the last part. After each part, pixel-level additions are used as identity connections.

reuse in the forward process of the network and gradient propagation in the backward process to improve the recovery of finer information.

Therefore, the dense blocks can link each layer in the network to share the information of the images, which helps to better train the semantics of the original image. Generally, the dense blocks contain an activation layer, a convolution layer, and a batch normalization layer. Since BN layer reduces the consistency of the removed image color and creates some artifacts [31]. To better remove the semantic information obscured by clouds using only the surrounding information, we extract and utilize the correlation functions in the convolution and activation layers of the residual module for correcting the semantics in training.

In Fig. 3, the dense block consists of five residual blocks. Each small residual block contains an activation layer and a convolutional layer. Except for the last residual block, the residual block is constructed by a stacked convolution and an ELU activation function. Each residual block is connected by the elementwise addition.

Dense block mainly combines the results of the last layer and this layer as the input of the next layer.  $X_i$  is the network input of the  $i$ th layer, then the output of the dense-connected block is expressed as

$$y_i = H_i \{x_0, x_{i-1}\} \quad (2)$$

where  $H_i$  is the nonlinear map of the  $i$ th layer, and  $x_0, x_1, \dots, x_i$  are the feature maps merged into the output of the 0th, 1st, and  $i$ th layers. The dense block merges the feature maps of each layer by introducing dense connections. It reduces parametric features and helps shallow features communicate effectively and reduce the loss of information in intermediate layers.

3) *Contextual Attention Module*: After the coarse removal, we obtained the general features of the initial removal image. Refinement based on coarse removal area to remove as much semantics as possible from the cloud-obscured area. In the refinement removal part, we adopt the contextual attention module.

As convolutional networks can only process features of local images layer by layer, it is difficult to effectively use feature information from far away in space for restoration. Therefore, we adopt the contextual attention module in the refinement removal network. It is used to recover the area where the thick clouds is located by continuously computing the matching scores between the convolutional background block (the know region) and the foreground (the area where the cloud is located) block, and selecting the optimal matching score as the semantics of the foreground area. The specific implementation of the contextual

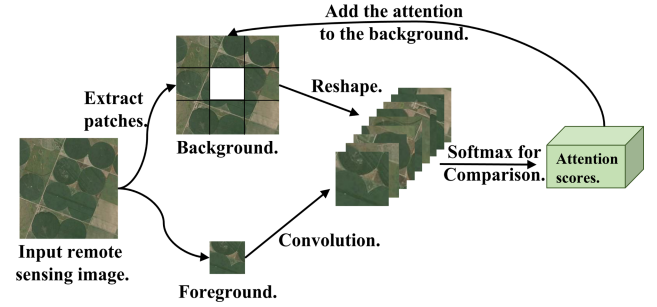


Fig. 4. Contextual attention module. It uses convolution to calculate the similarity between background areas and foreground (cloud areas), apply softmax to compare and get the attention score of each pixel. Foreground (cloud areas) with background semantics is reconstructed by deconvolution of the attention scores.

attention model is shown in Fig. 4, and the process can be concluded as follows: First, the convolution is used to match a similar area from the known image contents. Then, the softmax is done on the full channel to find the area that most like the areas to be repaired. Finally, the information of the areas is deconvoluted to reconstruct the patched areas.

In addition, the core of contextual attention is to focus on match and propagation. Contextual attention mainly matches the foreground features of the missing pixel with the background environment. For example, we extract patches (such as  $3 \times 3$ ) from the background as a new convolution filter and define patches of the foreground as  $f_{x,y}$  and the patches of background as  $b_{x',y'}$ . The similarity of patches centered on foreground and background is defined as  $S_{(x,y),(x',y')}$ , which is expressed as

$$S_{(x,y),(x',y')} = \left\langle \frac{f_{x,y}}{\|f_{x,y}\|}, \frac{b_{x',y'}}{\|b_{x',y'}\|} \right\rangle. \quad (3)$$

In order to make attention more consistent, the foreground patch should shift the same as the background patch. The contextual attention module takes the identity matrix as the core, and realizes the propagation of convolution more effectively. We first make the propagation with left-to-right and then top-to-down with the kernel size  $k$ , and the formula is

$$\hat{S}_{(x,y),(x',y')} = \sum_{i \in -k, \dots, k} S_{(x,y),(x',y')}^* \quad (4)$$

4) *Loss Function*: There are two main losses of DANet, one is the spatial reconstruction loss, and the other is the WGAN-GP loss.

We use spatial reconstruction loss to reduce the large ambiguity caused by the removed surface information. The spatial decay reconstruction loss is used to reduce the weight of the central area as the pixel points get closer to the central location so that the calculation of the loss value will not mislead the training process because the gap between the central result and the original image is too large. This part changes the weight into a distance-based weight  $\gamma^l$ , where  $l$  indicates the distance between the missing area and the closest known area. Obviously, the smallest known area is closest to the unknown area, when  $l$  is small, the value of  $\gamma^l$  is large, and the correlation is high. According to the empirical value  $\gamma$  is set to 0.99 [32].

For another loss, we use global and local WGAN-GP loss. WGAN-GP provides global and local outputs in the refinement network to enforce global and local consistency. Among them, one of the discriminators discriminates the whole image in a global sense, while the other discriminates the local images around the filled area in a local sense. The input of the global loss is the whole image, and the input of the local loss is the mask local image.

For WGAN-GP, the weight of  $(1 - M)$  is added to allow the discriminator to focus on identifying the missing parts, where  $M$  is the mask. WGAN-GP uses the distance with *Earth-Mover* distance  $W(I_r, I_g)$  to compare the distributions with the generated and real data. They mainly multiply the gradient and the input mask  $M$ , and the formula is

$$\begin{aligned} Loss = & -E_{\hat{x} \sim \hat{I}_r} D(\hat{x}) + E_{\hat{x} \sim \hat{I}_g} D(\hat{x}) \\ & + \lambda E_{\hat{x} \sim \hat{I}_x} (\|\nabla_{\hat{x}} D(\hat{x}) \odot (1 - M)\|_2 - 1)^2 \end{aligned} \quad (5)$$

where  $\hat{x}$  is sampled from the shortest path between points in  $I_r$  and  $I_g$  distributions,  $D(x)$  is the discriminator, and  $\lambda$  is the gradient penalty coefficient.

### III. EXPERIMENTS

The complete cloud removal framework includes CLM and cloud removal model. In this article, it focuses on proposing the cloud removal model for DANet. Therefore, we mainly discuss the performance of DANet in the experiments. In this section, we describe the experimental setting and introduce the implementation detail. Besides, we compare ours with representative methods in synthetic and real cloud images in the thick cloud removal and discuss the results of them.

#### A. Experimental Settings

1) *Training and Testing Data*: The synthetic clouds are important for the training processing as the existing visible light data have a single type of ground. In fact, the surface information is rich and diverse, and our selected training set preferably covers various types of the surface, such as forests, cities, oceans, deserts, etc. Therefore, we use synthetic RGB-NIR remote images as the training and testing datasets.

Based on the atmospheric cloud model  $I_c = (1 - I_{gray}) \odot I_g + I_g$  [22], [23], NMPU-RESISC45 [20] and NOAA [21] datasets are synthesized for RGB-NIR multispectral remote sensing images.  $I_{gray}$  is the gray thickness image. The cloud layer synthesis process is shown in Fig. 5.

For the synthetic cloud removal, all the test synthetic cloud textures cover at least 30% areas of the whole image. Among the synthetic multispectral remote image datasets, one of them is NMPU-RESISC45 [20]. It extracts three bands including 6, 5, and 4 in multispectral remote sensing images, which show natural (false color) images. The dataset is a public benchmark for the satellite image scene classification created by Northwestern Polytechnical University. It contains a wide variety kind of images, with a total of 35 000 remote sensing images. Besides, we remove the cloud classification in NMPU-RESISC45 from the training set to eliminate the influence of this classification

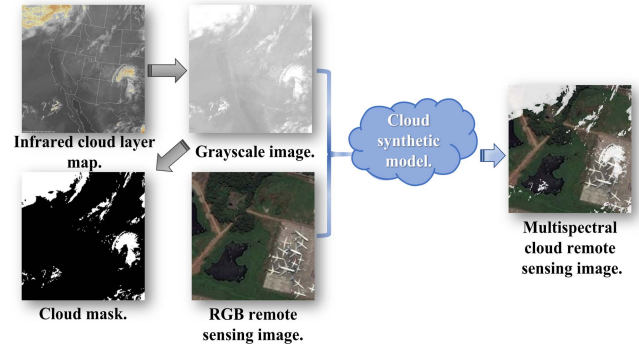


Fig. 5. Synthetic multispectral remote sensing. We first transform the infrared cloud layer map into a grayscale image. Then, the grayscale image and RGB remote sensing image are superimposed to synthesize the multichannel remote sensing image with clouds according to cloud synthetic model.

on the results. Other synthetic datasets are National Oceanic and Atmospheric Administration (NOAA) [21] and making the grayscale infrared image as the cloud images. Among them, 24 500 images of the dataset are applied for training, and the rest of them are used to create synthetic cloudy images as the testing dataset.

For the real cloud removal, the real cloud images are selected in the cloud classification of NMPU-RESISC45 datasets. Among them, most of the cloud cover more than 10% of the whole image, and we randomly select six images for discussion.

In this experiment, we mainly test the performance of our proposed algorithm by simulating the synthetic and real cloud images removal.

2) *Comparative Methods*: To verify the performance of our proposed algorithm, we compare it with state-of-the-art single remote sensing image cloud removal algorithms. We select six representative algorithms from traditional algorithms and deep learning algorithms. Since our proposed DANet is based on deep learning to remove thick clouds from remote sensing images, in the comparison of experimental algorithms, we keep a traditional algorithm low rank matrix completion (LRMC) [33] in 2016, which achieve better removal results. Deep learning algorithms include the gated convolution (GC) [34] proposed in 2019 and Zheng et al. adopted GAN in the remote sensing image to remove the cloud (short as GAN) [16] to remove the thick clouds from remote sensing images. Zhang et al. proposed STS-CNN [17], a remote sensing image reconstruction method that can be based on deep convolutional neural network in 2018, and proposed another network based on deep spatiotemporal prior combined with low-rank tensor singular value decomposition (DP-LRTSVD) in 2021 [18], and Xu et al. proposed Landsat image decoupling generation adversarial network (AMGAN-CR) method based on attention mechanism in 2022 [19].

3) *Parameter Settings*: The setting of propagating kernel  $k$  [Eq. (4)] affects inpainting effect. For the setting of  $k$ , the minimum patch is set to  $3 \times 3$ , and the maximum patch is set to  $7 \times 7$ . Therefore, we discuss the settings of the kernel size of  $k$ :  $3 \times 3$ ,  $5 \times 5$ , and  $7 \times 7$ . We test on 150 remote sensing images in the testing set and average their PSNR, RMSE, and

TABLE I  
AVERAGE VALUE OF THE TESTING SET OF PSNR, RMSE, AND SSIM WITH  
DIFFERENT KERNEL SIZE OF  $k$

Values for $k$	RMSE	PSNR	SSIM
$3 \times 3$	<b>0.0124</b>	<b>29.6930</b>	<b>0.9230</b>
$5 \times 5$	0.0127	29.493	0.9211
$7 \times 7$	0.0128	29.079	0.9204

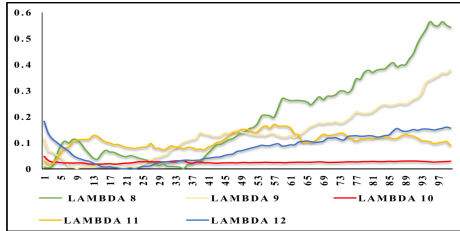


Fig. 6. Negative critic loss curve with several values of  $\lambda$ .

SSIM values, as shown in Table I. In Table I, when  $k$  is  $3 \times 3$ , the RMSE value is very small, PSNR and SSIM values are large, indicating that the best effect is obtained when this value is selected. When  $k$  is  $7 \times 7$ , the RMSE value is very large, PSNR and SSIM values are small, indicating that this value is the worst effect. Therefore, as the value of  $k$  increases, the repair effect becomes worse, and we set the value of  $k$  to  $3 \times 3$ .

For the setting of gradient penalty coefficient  $\lambda$  of Eq. (5), this parameter affects the convergence curve and speed of loss. In order to verify the optimal of  $\lambda$ , we set the parameter from 8 to 12, and then we observe its negative critic loss convergence curve. As shown in Fig. 6, the negative critic loss curve of them is 100 epochs in training. We can see that when  $\lambda$  is 10, it has converged to the 40<sup>th</sup> epoch, and the fluctuation is not too large. The other epochs have not converged to 100 epochs. Therefore, the value of  $\lambda$  is also set to 10 in the experiments.

In the training settings, the batch size is set to 16. The DANet is implemented on TensorFlow v1.12.0, CUDNN v7.5, CUDA v10.0, and the used GPU is Nvidia Tesla v100.

4) *Evaluation Standards*: In the synthetic remote sensing images cloud removal, we use the cloud-free image as a reference, and then compare the visual effect of the image after cloud removal and whether the semantics is close to the cloud-free image to judge good or bad removal effect. Moreover, we evaluate the effect of declouding by calculating the PSNR, SSIM, and RMSE values of the declouded image and the original image. Among them, PSNR is the peak signal-to-noise ratio, which is the most common measurement tool to evaluate the image quality, and the larger the PSNR value, the better the cloud removal effect. The structural similarity metric (SSIM) is a metric used to evaluate the similarity between two images and judge the image quality considering the degradation of structural information, the larger the SSIM value the better the cloud removal effect. The RMSE is the calculation of the root mean square error between the real image and the corrected image, the smaller the RMSE value the better the cloud removal effect.

In real cloud removal, since there is no cloud-free as a reference, we can only evaluate from visual semantics. We mainly

observe whether the semantic of the region after cloud removal is logical with the surrounding region, whether the color of the region after cloud removal matches with the surrounding, and the semantic fluency, etc.

## B. Cloud Removal Result

1) *Synthetic Cloud Removal*: In order to verify the cloud removal performance of our algorithm, we simulate the coverage of cloud from 5% to 50%. Generally, it is difficult to remove the original ground information when the cloud amount exceeds 50%. We add comparative experiments and count their removal parameter in Fig. 7.

Fig. 7 show the line charts of PSNR, RMSE, and SSIM with different cloud coverage. In Fig. 7, with the increase of cloud cover, the repairing effect performance of all algorithms gradually deteriorates. Compared with other algorithms, the proposed algorithm can maintain the optimal performance under various cloud amounts, especially for the large-size cloud removal.

Besides, we randomly select six representative images and display them in Fig. 8 (cloud coverage more than 30%). For the cloud removal result, we analyze the effects of semantic restoration, whether there are artifacts and noise. The highlighted areas with a red circle demonstrate the difference between the cloud removal algorithms. The first row of Fig. 8 shows the Ground image, the second row shows the synthetic multispectral cloud image, and rows 3–9 show the cloud removal results of LRMC [33], STS-CNN [17], GC [34], DP-LRTSVD [18], GAN [16], AMGAN-CR [19], and our algorithm, respectively.

In image 1, the semantics of the road after LRMC removal in the red circle in the upper-left of the first image is more dependent on the semantics of the surrounding environment, and the AMGAN-CR and ours can recover the original semantics of the road better. For the house semantic recovery, GC, AMGAN-CR and ours can recover the shape of the house, and our algorithm can recover the texture of the roof more closely while the house boundary is blurred after DP-LRTSVD and GAN restoration. In image 2, LRMC and our algorithm recover rounder and the recovered by our algorithm are closer to the original image in the upper-right red circle, but the semantics recovered by other algorithms have some deviations. For bottom-left red circle, GC and our algorithm are able to recover the semantics of the road. In image 3, the semantics of the original road is disconnected in the left circle, only GAN and our algorithm are able to recover it closer. In the red circle on the right side, the original semantics is an oval open space. LRMC, AMGAN-CR, and our algorithm can recover the original appearance of the road, but the road recovered by LRMC has a little distortion, while the semantics of the road recovered by GAN and CG has some deviation. In image 4, the original semantics in the red circle is a road, DP-LRTSVD, GAN, AMGAN-CR, and our proposal can reasonably restore the road in the red circle, and the road recovered by our algorithm more smoothly, while LRMC restores the road and generates redundant semantics, STS-CNN restores it fuzzier. In the bottom-right red circle, only GC and our algorithm recover closer to the semantics of the original image, i.e., the open space and the road are separated from each other. In image

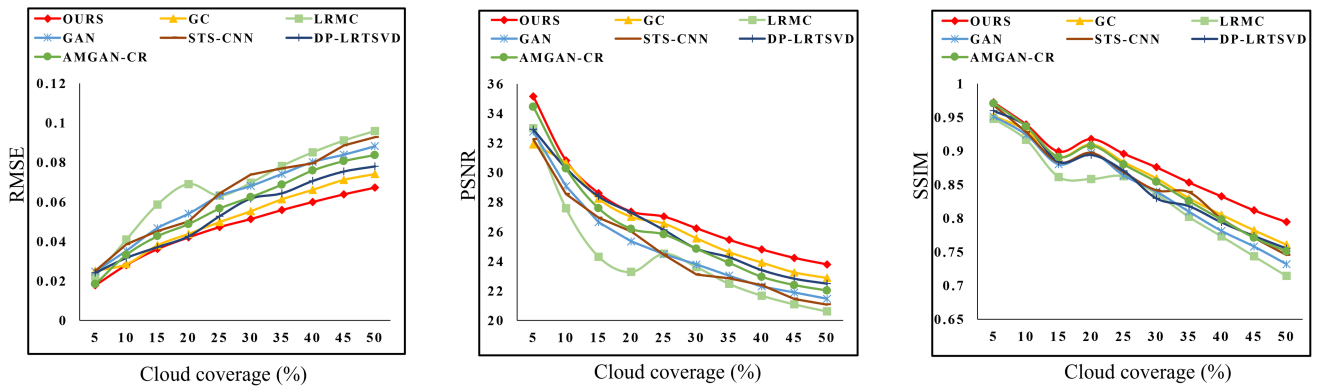


Fig. 7. Line chart comparison of the PSNR, RMSE, and SSIM after cloud removal with several algorithm from 5% to 50% cloud coverage.

TABLE II  
RESULTS OF SEVERAL DIFFERENT METHODS ON PSNR, RMSE, AND SSIM VALUES OF SYNTHETIC REMOTE SENSING IMAGES AFTER CLOUD REMOVAL

Images	Performance Parameters	Methods						
		LRMC [33]	STS-CNN [17]	GC [34]	DP-LRTSVD [18]	GAN-base [16]	AMGAN-CR [19]	Ours
Image 1	PSNR	23.4883	24.6066	25.6297	24.8504	23.9711	25.7709	<b>26.0433</b>
	RMSE	0.0197	0.0184	0.0177	0.0182	0.0184	0.0174	<b>0.0173</b>
	SSIM	0.8715	0.8906	0.8987	0.8703	0.8852	0.8945	<b>0.9067</b>
Image 2	PSNR	29.0083	27.6770	28.3532	27.3719	29.7821	29.7764	<b>30.1918</b>
	RMSE	0.0150	0.0178	0.0156	0.0179	0.0138	0.0148	<b>0.0133</b>
	SSIM	0.9480	0.9316	0.9365	0.9220	0.9517	0.9541	<b>0.9568</b>
Image 3	PSNR	27.6544	25.5782	25.2458	25.3963	28.3141	27.8081	<b>29.0610</b>
	RMSE	0.0178	0.0187	0.0204	0.0203	0.0160	0.0172	<b>0.0159</b>
	SSIM	0.8863	0.8659	0.8572	0.8409	0.8898	0.8939	<b>0.9107</b>
Image 4	PSNR	23.4551	24.0877	23.4938	24.7248	25.7415	24.1867	<b>26.4022</b>
	RMSE	0.0199	0.0185	0.0190	0.0170	0.0154	0.0177	<b>0.0149</b>
	SSIM	0.9056	0.9147	0.8976	0.9110	0.9144	0.9256	<b>0.9349</b>
Image 5	PSNR	28.1043	28.1787	32.3120	29.7011	29.1337	32.0477	<b>32.3226</b>
	RMSE	0.0191	0.0183	0.0132	0.0144	0.0158	0.0136	<b>0.0125</b>
	SSIM	0.9055	0.9124	0.9508	0.9107	0.9210	0.9470	<b>0.9767</b>
Image 6	PSNR	21.3296	20.6421	24.4656	21.0136	23.5593	22.0810	<b>24.9879</b>
	RMSE	0.0227	0.0261	0.0204	0.0240	0.0206	0.0207	<b>0.0201</b>
	SSIM	0.8892	0.8721	0.9176	0.8613	0.9091	0.9008	<b>0.9267</b>
The average value of testing set	PSNR	28.0881	27.9600	27.6501	27.4875	29.0074	28.8451	<b>29.6930</b>
	RMSE	0.0139	0.0151	0.0163	0.0174	0.0134	0.0135	<b>0.0124</b>
	SSIM	0.8937	0.9120	0.8937	0.8794	0.9206	0.9159	<b>0.9230</b>

5, only CG, AMGAN-CR, and our proposal can reasonably restore the river in the red circle, and the roads recovered by AMGAN-CR and our algorithm are smoother. In image 6, except DP-LRTSVD method, all other methods can basically restore the basic boundary of the lake, and the boundary of the lake restored by AMGAN-CR and our method is more natural.

In general, for large area cloud removal, most algorithms can remove it relatively well when the image semantics is relatively simple and regular. In complex semantic images, our proposed can better remove the original semantics before cloud occlusion. For LRMC algorithm, it can repair images with simple regular semantics, and the repair effect is not satisfactory in complex semantics. The semantics is severely distorted after cloud removal by the STS-CNN algorithm. The GC and DP-LRTSVD

algorithms are prone to some artifacts after removal of the cloud. The semantics of GAN is not particularly natural in the cloud removal. AMGAN-CR does not have artifacts after restoration, but in terms of semantics, our proposed algorithm restores the semantic information closer to the original image. Therefore, the advantage of our algorithm is that when the cloud amount accounts for more than 30% of the hole image, our algorithm can achieve better cloud removal effect than other algorithms and can be closer to the original cloud-free image.

Table II shows the quantitative results with six images on PSNR, SSIM, and RMSE. We take the ground image as the reference standard image and calculate the gap between each algorithm after cloud removal and the original image, the smaller the

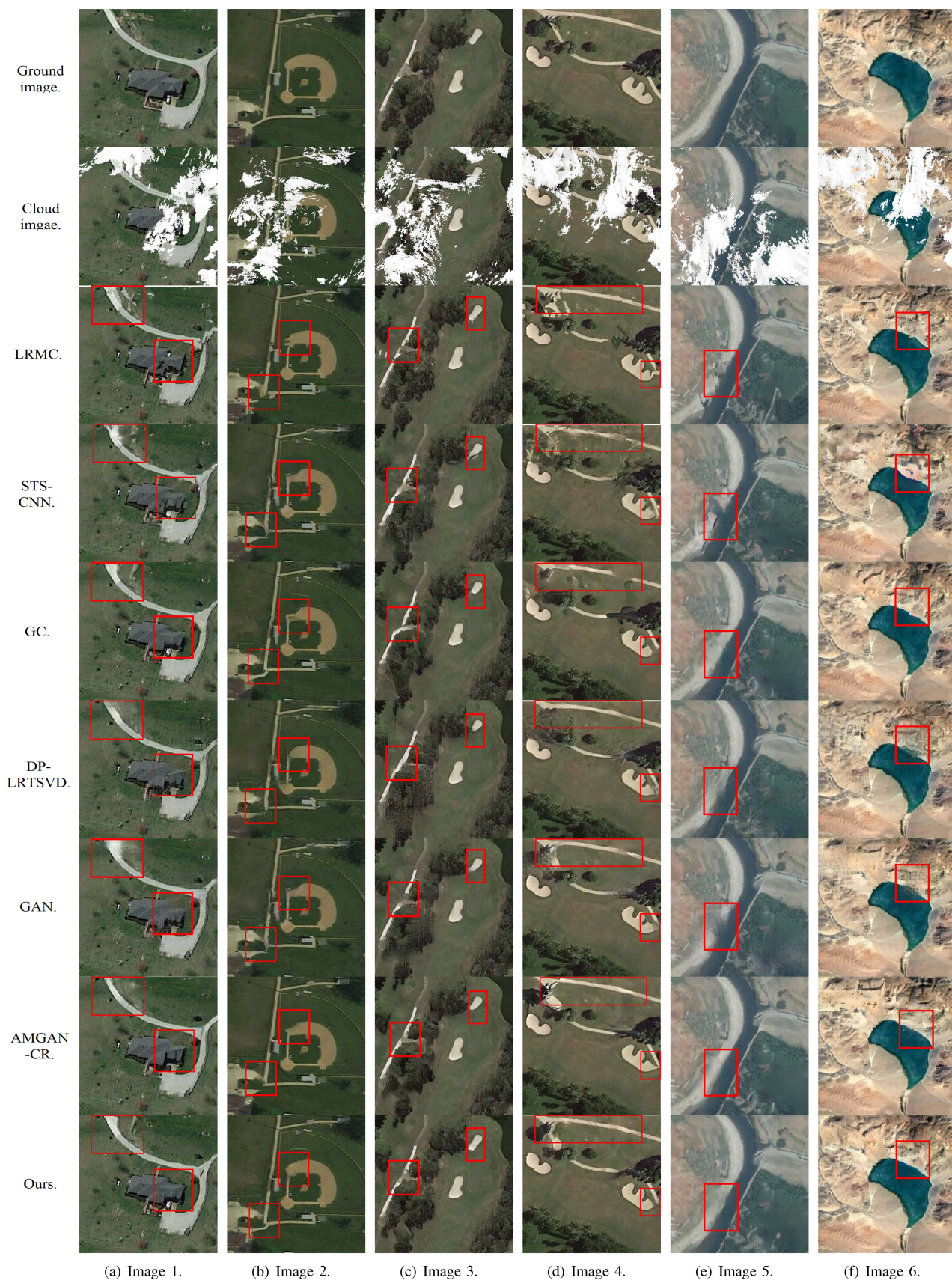


Fig. 8. Comparison the cloud removal results of synthetic remote sensing images with the different methods. The cloud coverage of the whole images are more than 30%. The red area mask is the difference between the cloud removal with these methods. Our algorithm can better remove the original semantics before cloud occlusion and without redundant artifacts.



gap indicates the better effect of cloud removal. Our algorithm achieves the best values in terms of PSNR, SSIM, and RMSE of the six images. The average PSNR value on the test dataset is higher than that of LRMC, STS-CNN, GC, DP-LRTSVD, GAN-base, and AMGAN-CR, respectively: 1.6049 dB, 1.7330 dB, 2.0429 dB, 2.2065 dB, 0.6856 dB, and 0.8479 dB. Ours calculation is 0.034 higher than the second highest algorithm GAN and 0.0436 higher than the lowest algorithm DP-LRTSVD in SSIM. It appears that our algorithm has better performance.

2) *Real Cloud Removal*: In the real image cloud removal, we select six images with obvious thick clouds for analysis. Fig. 9 shows the real cloud removal result. The first row of Fig. 9 shows the cloud image, the second row shows cloud mask image extracted by CLM, and rows 3–9 show the cloud removal results of LRMC [33], STS-CNN [17], GC [34], DP-LRTSVD [18], GAN [16], AMGAN-CR [19], and our algorithm, respectively.

In the real cloud removal of a single image, since there is no cloud-free image as a reference for the cloud removal effect, we analyze the performance of each algorithm from the perspective of visual semantics.

In the image 1, the clouds obscured are part of the mountains, where LRMC and GAN, DP-LRTSVDAMGAN-CR have obvious recovery traces, the semantic recovered by GC is less consistent around, and STS-CNN recovered semantics have artifacts present, the semantics recovered by our proposed algorithm is more natural with the surrounding. In the image 2, our algorithm is able to recover the semantics that no longer produce new clouds after recovery. Although the color is closer to the surrounding mountains after LRMC and CG recovery, it is easy to produce new clouds after cloud removal like other algorithms. In the cloud removal of the image 3, we only remove the thick cloud part. There is some visual noise after GC cloud removal, while the color after LRMC and GAN cloud removal is not consistent with the surrounding environment. Besides, there is some color deviation after GC and AMGAN-CR cloud removal, and other methods can achieve better results. In the image 4, there is a certain amount of noise in LRMC and GC, and the color after STS-CNN and AMGAN-CR recovery is white biased. In the image 5, the semantic restoration of all algorithms is not satisfactory. In the image 6, the STS-CNN recovery produces artifacts, the GC, and AMGAN-CR recovered colors are less unloaded from the surrounding. Other algorithms are able to recover more naturally. In image 1 to 3, our algorithm removal the cloud is more logical in visual semantics, and the colors are more consistent with the surroundings. However, in image 5, the semantic recovery of all algorithms is not satisfactory because the cloud-obscured area is located at the intersection of a river and a field, and the semantics are relatively complex. In contrast, ours and GAN can remove part of the river area after cloud removal, and the color is more coherent with the surrounding environment.

In general, our algorithm is closer to the surrounding semantic state for real cloud removal. The removal result with DANet is relatively smoother than others in the terms of semantic.

TABLE III  
AVERAGE PERFORMANCE OF PSNR, RMSE, AND SSIM OF REPLACING WITH CNN, ADDING DENSE BLOCK, ADDING CA MODEL, AND OUR PROPOSED METHOD

Performance Parameters	Replacing with CNN	Adding dense block	Adding CA Moudle	Ours
PSNR	27.9600	28.6610	29.0936	<b>29.6930</b>
RMSE	0.0389	0.0133	0.0151	<b>0.0124</b>
SSIM	0.9120	0.9140	0.9137	<b>0.9230</b>

## IV. DISCUSSION

### A. Ablation Studies

In DANet, we design a dense block in the coarse network based on CNN for training to recover features from weak boundary information in each pixel direction. In refinement of the network, to improve the semantics of regions obscured by clouds, we construct the contextual attention module. Therefore, in the ablation study, in order to verify whether the dense block and contextual attention module we designed are important in DANet, we conducted experiments in the ablation experiment without adding the dense block and CA module, adding only the dense block and adding only the contextual attention module, respectively.

In Fig. 10(c), it can be seen that if there is no dense block and CA module but only a simple CNN network, the effect of cloud removal is not satisfactory and artifacts are easily generated. Only adding the dense block [see Fig. 10(d)] can effectively solve the problem of artifacts, but the semantics of the cloud removal has some deviation from the surrounding environment. Only adding CA module [see Fig. 10(e)] without dense block can recover the original semantic information of the ground to a greater extent, but it is easy to generate artifacts. Only when the dense block and CA module [see Fig. 10(f)] are added simultaneously, the effect of the cloud removal is closest to restoring the original image semantics of the ground, and there are no unnecessary artifacts.

In the numerical statistics of the test set as shown in Table III, we can see that the cloud removal values can outperform the CNN network regardless of whether the dense block or CA module is added, and the best cloud removal results can be achieved when both the dense block and CA module are added.

In summary, CA module can effectively improve the semantics of cloud removal, and dense block can effectively remove artifacts to obtain the best cloud removal effect.

### B. Robustness and Generalizability

In order to discuss the robustness and generalizability of our algorithm, we verify it from other complex scenes.

In this experiment, we select four remote sensing images with complex background semantics and randomly selected four cloud images for synthesis, and the result shows in Fig. 11. In the upper left corner of the third row of the images, the semantic information is recovered differently from the original image because the ground information is blocked by a large number of clouds. However, the semantic vision after cloud removal

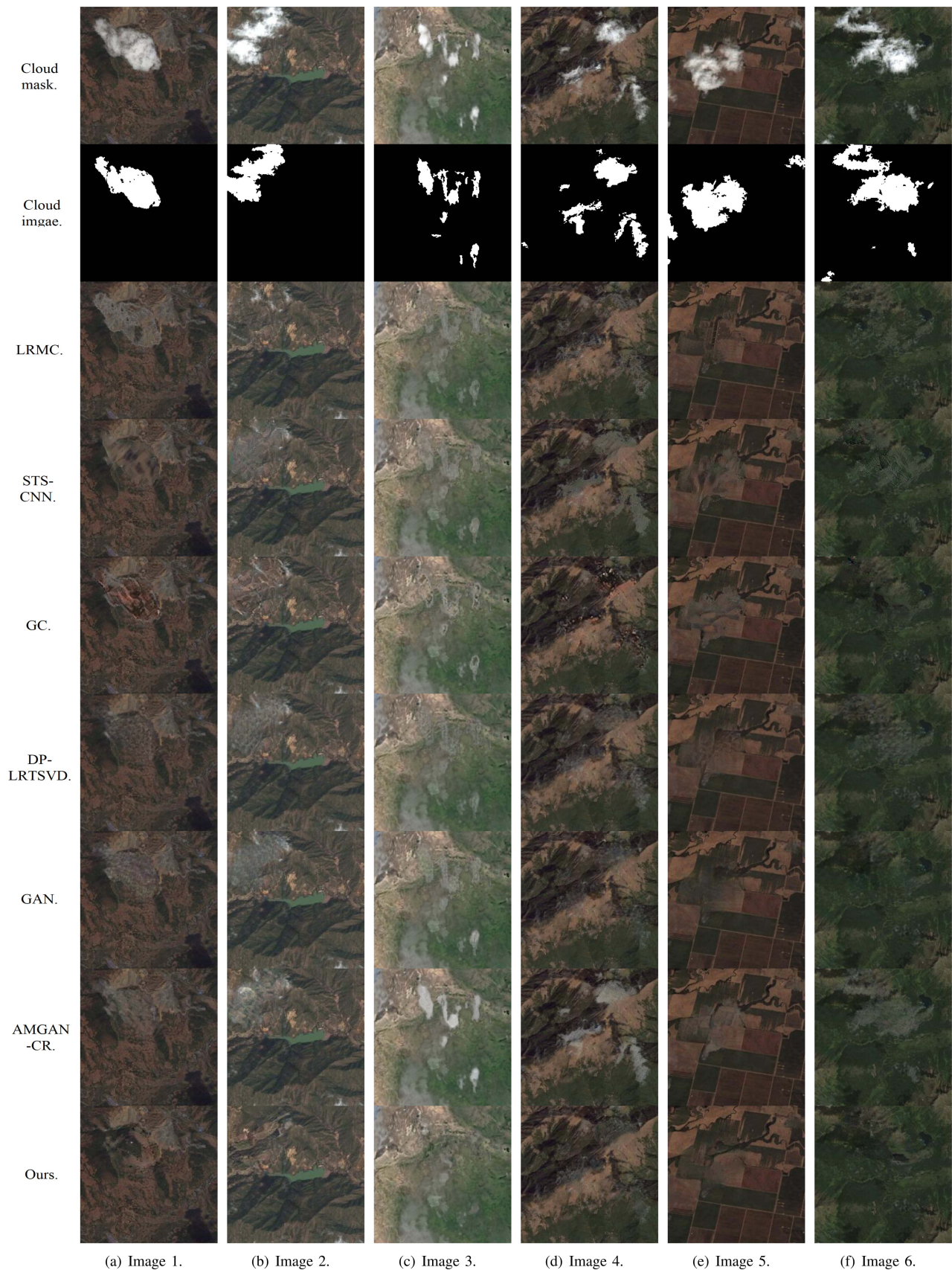


Fig. 9. Comparison the cloud removal results of real remote sensing images with the different methods. Our algorithm is closer to the surrounding semantic state and is relatively smoother than other algorithms in the terms of semantic.

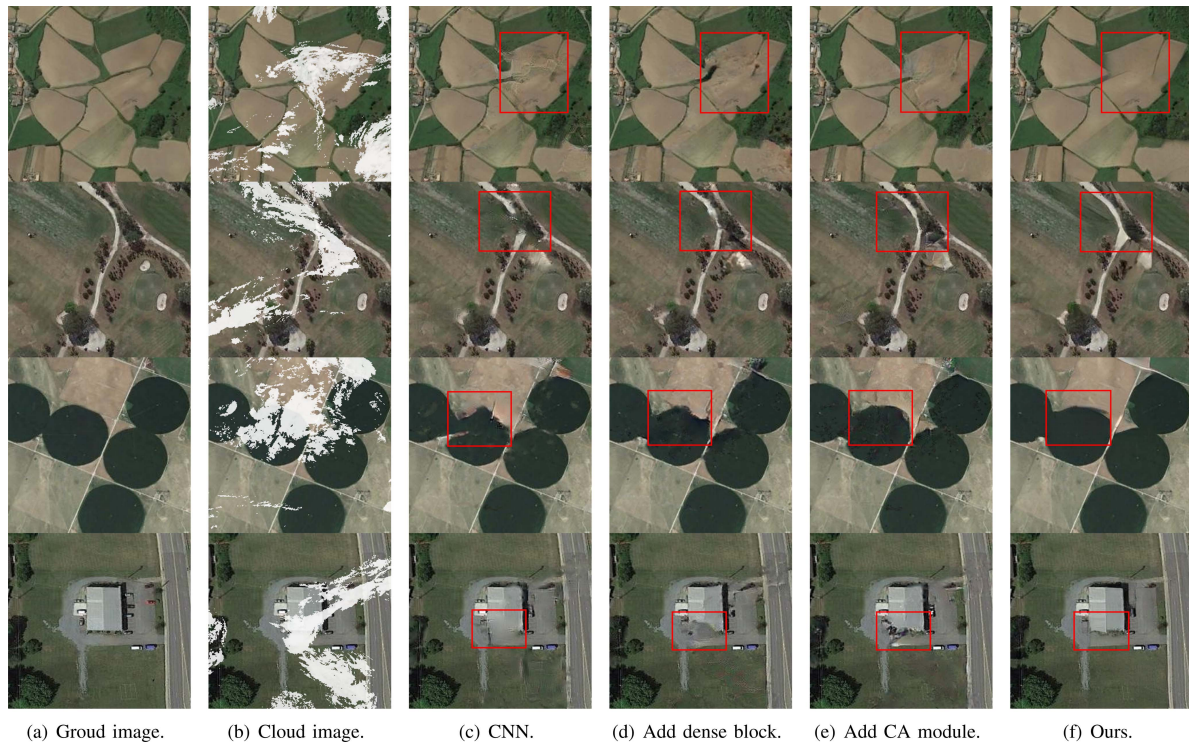


Fig. 10. Ablation study. (a) Ground images. (b) Cloud image. The red region mask is the difference between the cloud removal with (c) CNN (remove the dense block and contextual attention module), (d) only add the dense block (without CA module), (e) only add the CA module (without the dense block), and (f) ours.

TABLE IV  
RESULTS OF SEVERAL DIFFERENT METHODS ON PSNR, RMSE, AND SSIM VALUES OF SYNTHETIC REMOTE SENSING IMAGES AFTER CLOUD REMOVAL

Performance Parameters	Methods						
	LRMC [33]	STS-CNN [17]	GC [34]	DP-LRTSVD [18]	GAN-base [16]	AMGAN-CR [19]	Ours
Training time (h)	--	3.5	65	74	34	5	58
Test time (s)	2296	0.3692	0.2924	0.5147	0.7372	0.3782	0.2661



Fig. 11. Thick cloud removal in other complex scenes. The first row is the raw image, the second row is the synthetic cloud image, and the third row is the image after cloud removal.

still conforms to the structural logic and the edges are relatively coherent. In other images, the effect after removing the cloud is closer to the original image when the area occluded by the cloud is not particularly large.

Therefore, in the cloud removal of complex scenes, DANet can more accurately remove the original semantic information of the ground when the edge semantic information is still retained after the cloud layer is occluded.

### C. Complexity Analysis

We analyze the model time cost of DANet while comparing it with the time cost of other comparative experiments. From Table IV, we can see that the training time of our algorithm takes 58 h. The training time of STS-CNN algorithm takes at least 3.5 h, but it has the worst cloud removal effect. The longest training time is 74 h for DP-LRTSVD algorithm. In the test time, the traditional algorithm LRMC takes the longest time of 2296 s, compared to the deep learning algorithm test time is less, all within 1 s. Our algorithm test time is at least 0.2661 s. Although we do not have much advantage in the training time, we have the shortest average testing time. As long as the same type of dataset is available, we only need to train once, and the later ones can be tested directly. Moreover, we can see in the previous comparison experiments that our algorithm can work better than other deep learning algorithms after large area cloud removal.

## V. CONCLUSION

In this article, the thick cloud removal framework by reconstructing semantic information based on a single remote sensing image is designed, and it can recover the ground information occluded by cloud as much as possible by using the limited boundary information. First, the cloud layer is located and extra by the CLM that facilitate the thick cloud removal. Then, a cloud removal network (DANet) is designed, which performs recovery of the areas obscured by thick clouds. To better recover the semantic information of the cloud-occluded area, we design a dense block and a contextual attention module in DANet. The experimental results show that the proposed method can recover the original semantic information on the ground more accurately, while using more limited boundary information than some state-of-the-art cloud removal methods.

In the future, we will make our algorithm more useful in remote sensing image collection and application.

## REFERENCES

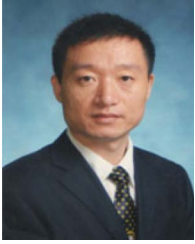
- [1] C.-H. Lin, P.-H. Tsai, K.-H. Lai, and J.-Y. Chen, "Cloud removal from multitemporal satellite images using information cloning," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 232–241, Jan. 2013.
- [2] J. Ju and D. P. Roy, "The availability of cloud-free landsat ETM data over the conterminous United States and globally," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 1196–1211, 2008.
- [3] R. Mao, H. Li, G. Ren, and Z. Yin, "Cloud removal based on SAR-optical remote sensing data fusion via a two-flow network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 7677–7686, 2022.
- [4] Y. Chen, L. Tang, X. Yang, R. Fan, M. Bilal, and Q. Li, "T images using deep learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 143–153, 2020.
- [5] Y. Zhang, F. Wen, Z. Gao, and X. Ling, "A coarse-to-fine framework for cloud removal in remote sensing image sequence," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5963–5974, Aug. 2019.
- [6] C. Tao, S. Fu, J. Qi, and H. Li, "Thick cloud removal in optical remote sensing images using a texture complexity guided self-paced learning method," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–12, 2022.
- [7] M. Xu, X. Jia, M. Pickering, and A. J. Plaza, "Cloud removal based on sparse representation via multitemporal dictionary learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2998–3006, May 2016.
- [8] C.-H. Lin, K.-H. Lai, Z.-B. Chen, and J.-Y. Chen, "Patch-based information reconstruction of cloud-occluded multitemporal images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 163–174, Jan. 2014.
- [9] H. Shen et al., "Missing information reconstruction of remote sensing data: A technical review," *IEEE Trans. Geosci. Remote Sens.*, vol. 3, no. 3, pp. 61–85, Sep. 2015.
- [10] R. Cao, Y. Chen, J. Chen, X. Zhu, and M. Shen, "Thick cloud removal in landsat images based on autoregression of landsat time-series data," *Remote Sens. Environ.*, vol. 249, 2020, Art. no. 112001.
- [11] R. E. Rossi, J. L. Dungan, and L. R. Beck, "Kriging in the shadows: Geostatistical interpolation for remote sensing," *Remote Sens. Environ.*, vol. 49, no. 1, pp. 32–40, 1994.
- [12] F. Van der Meer, "Remote-sensing image analysis and geostatistics," *Int. J. Remote Sens.*, vol. 33, no. 18, pp. 5644–5676, 2012.
- [13] C. Zhang, W. Li, and D. Travis, "Gaps-fill of SLC-off landsat ETM satellite image using a geostatistical approach," *Int. J. Remote Sens.*, vol. 28, no. 22, pp. 5103–5122, 2007.
- [14] A. C. Siravenha, D. Sousa, A. Bispo, and E. Pelaes, "Evaluating inpainting methods to the satellite images clouds and shadows removing," in *Proc. Int. Conf. Signal Process., Image Process., Pattern Recognit.*, 2011, pp. 56–65.
- [15] L. Lorenzi, F. Melgani, and G. Mercier, "Multiresolution inpainting for reconstruction of missing data in VHR images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2011, pp. 531–534.
- [16] J. Zheng, X.-Y. Liu, and X. Wang, "Single image cloud removal using u-net and generative adversarial networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6371–6385, Aug. 2021.
- [17] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial-temporal-spectral deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4274–4288, Aug. 2018.
- [18] Q. Zhang, Q. Yuan, Z. Li, F. Sun, and L. Zhang, "Combined deep prior with low-rank tensor SVD for thick cloud removal in multitemporal images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 177, pp. 161–173, 2021.
- [19] M. Xu, F. Deng, S. Jia, X. Jia, and A. J. Plaza, "Attention mechanism-based generative adversarial networks for cloud removal in landsat images," *Remote Sens. Environ.*, vol. 271, 2022, Art. no. 112902.
- [20] G. Cheng, J. Han, and X. Lu, "Remote sensing image scene classification: Benchmark and state of the art," *Proc. IEEE*, vol. 105, no. 10, pp. 1865–1883, Oct. 2017.
- [21] NESDIS, "Colorized infrared from american national oceanic and atmospheric administration (NOAA)," Accessed: Nov. 2020. [Online]. Available: <https://www.nesdis.noaa.gov/content/colorized-infrared>
- [22] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.
- [23] R. Fattal, "Single image dehazing," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–9, 2008.
- [24] 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.
- [25] J. Yin, T. Wang, Y. Du, X. Liu, L. Zhou, and J. Yang, "Slic superpixel segmentation for polarimetric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–17, 2022.
- [26] C. Tian, Y. Xu, W. Zuo, B. Zhang, L. Fei, and C.-W. Lin, "Coarse-to-fine CNN for image super-resolution," *IEEE Trans. Multimedia*, vol. 23, pp. 1489–1502, 2021.
- [27] X. Wang et al., "ESRGAN: Enhanced super-resolution generative adversarial networks," in *Proc. Eur. Conf. Comp. Vis. Workshops (ECCVW)*, Sep. 2018, pp. 1–16.
- [28] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn., Ser. Proc. Mach. Learn.*, F. Research Bach and D. Blei, Eds., 2015, vol. 37, pp. 448–456.
- [29] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," 2015, *arXiv:1511.07289*.
- [30] W. Yang, X. Li, and L. Zhang, "Toward semantic image inpainting: Where global context meets local geometry," *J. Electron Imag.*, vol. 30, pp. 023028–023028, 2021.
- [31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.
- [32] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 5505–5514.
- [33] J. Wang, P. A. Olsen, A. R. Conn, and A. C. Lozano, "Removing clouds and recovering ground observations in satellite image sequences via temporally contiguous robust matrix completion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2754–2763.
- [34] D. Chen et al., "Gated context aggregation network for image dehazing and deraining," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2019, pp. 1375–1383.



**Yuyun Chen** received the B.S. degree in electrical engineering and automation engineering from the Beijing Institute of Technology, Zhuhai, China, in 2018, and the M.S. degree in computer and information systems from the School of Computer Science and Engineering, Macau University of Science and Technology, Taipa, Macau, China, in 2020. She is currently working toward the Ph.D. degree in computer technology and application with the School of Computer Science and Engineering, Faculty of Innovation Engineering, Macau University of Science and Technology, Taipa, Macau, China.

Her research interests include the areas of remote sensing image inpainting and deep learning.

Dr. Chen was a recipient of the First Prize at the 17th and 18th China Post-graduate Mathematical Contest in Modeling, and Second Prize at the 15th and 16th China Post-graduate Mathematical Contest in Modeling.



**Zhanchuan Cai** (Senior Member, IEEE) received the Ph.D. degree in computer software and theory from Sun Yat-sen University, Guangzhou, China, in 2007.

He is currently a Professor with the School of Computer Science and Engineering, Macau University of Science and Technology, Macau, China, where he is also with the State Key Laboratory of Lunar and Planetary Sciences at Macau University of Science and Technology. He has authored more than 100 papers in journals and refereed conference. His research interests include image processing and computer graphics,

intelligent information processing, multimedia information security, and remote sensing data processing and analysis.



**Jieyu Yuan** received the M.S. degree in computer and information systems from the School of Computer Science and Engineering, Macau University of Science and Technology, Taipa, Macau, China, in 2021. He is currently working toward the Ph.D. degree in computer technology and application from the School of Computer Science and Engineering, Faculty of Innovation Engineering, Macau University of Science and Technology, Taipa, Macau, China.

His research interests include remote sensing images processing, underwater images enhancement,

and machine learning.



**Lianghai Wu** received the B.S. degree in computer science from the School of Computer Science, South China Normal University, Guangzhou, China, in 2000, and the M.S. degree in software engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2006. He is currently working toward the Ph.D. degree in computer technology and application from the School of Computer Science and Engineering, Faculty of Innovation Engineering, Macau University of Science and Technology, Taipa, Macau, China.

His current research interests include remote sensing data processing and analysis.