






## Automatic Cloud Detection and Removal in Satellite Imagery Using Deep Learning Techniques



Jingyi Li , Yinbao Lv , Xu Yan , Hongjian Weng\* , Duo Li , Nan Shi 

Operations Services Department China Satellite Network Application Co., Ltd, Beijing 100000, China

Corresponding Author Email: [wenghj@pku.edu.cn](mailto:wenghj@pku.edu.cn)

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.410226>

### ABSTRACT

**Received:** 8 November 2023

**Revised:** 12 February 2024

**Accepted:** 27 February 2024

**Available online:** 30 April 2024

#### Keywords:

*satellite imagery, cloud detection, cloud removal, superpixel segmentation, generative adversarial networks (GAN), deep learning*

With the rapid advancement of remote sensing technology, satellite imagery has become increasingly vital in global geographic information systems, environmental monitoring, and resource management. However, cloud cover frequently degrades the quality of satellite images, limiting their effectiveness in many critical areas. Traditional methods for cloud detection and removal, such as threshold analysis and spectral feature analysis, often fail to achieve satisfactory results due to environmental constraints and algorithmic limitations. In response, this study employs deep learning techniques, specifically superpixel segmentation and generative adversarial networks (GAN), to address this issue. This paper begins by discussing the importance of cloud detection and removal in satellite imagery and reviews existing major techniques and methods. It then explores the application of superpixel segmentation based on local adaptive distance for automatic cloud boundary identification, along with innovative applications of GAN for surface information reconstruction in cloud-covered areas. These methods not only enhance the accuracy of cloud detection but also effectively optimize the cloud removal process, paving the way for further applications of satellite imagery.

## 1. INTRODUCTION

With the rapid development of remote sensing technology, satellite imagery has become an important tool for Earth observation and environmental monitoring [1-3]. However, cloud cover often severely affects the quality and utility of satellite images, limiting their effectiveness in fields such as meteorological analysis, resource surveys, and environmental protection [4-7]. Therefore, developing effective cloud detection and removal techniques is crucial for enhancing the usability and accuracy of satellite data.

Currently, deep learning-based techniques have demonstrated outstanding performance in the field of image processing, especially in image segmentation and feature recognition [8-10]. Applying these techniques to cloud detection and removal in satellite imagery can significantly improve processing precision and efficiency [11, 12]. Moreover, research on cloud detection and removal not only enhances the practicality of satellite imagery in real-world applications but may also promote the development of automatic processing and analysis technologies for remote sensing images.

Despite existing research providing various methods for cloud detection and removal, such as threshold-based detection and spectral feature analysis, these methods often rely on strict environmental assumptions or have high error rates [13-16]. Additionally, traditional methods frequently face issues of insufficient accuracy and low efficiency when dealing with complex or changing cloud structures [17-20]. Therefore, a more precise and adaptive technology is required

to cope with the variable effects of cloud cover.

This paper primarily studies two aspects. First, it employs a satellite image automatic cloud boundary-sensitive superpixel segmentation technique based on local adaptive distance. This technique, by analyzing the complex relationships between pixels, adaptively determines the cloud boundaries, thereby improving the accuracy of cloud detection. Second, it develops an automatic cloud removal algorithm for satellite images based on GAN, utilizing the powerful generative capabilities of GAN to restore surface information in cloud-covered areas. The combination of these two technologies not only brings new research perspectives to the field of cloud detection and removal but also greatly enhances the value and reliability of satellite images in various practical applications.

## 2. SUPERPIXEL SEGMENTATION FOR AUTOMATIC CLOUD BOUNDARY DETECTION IN SATELLITE IMAGERY BASED ON LOCAL ADAPTIVE DISTANCE

In satellite image analysis, the detection and removal of clouds is a crucial yet challenging task, especially when dealing with scenes where cloud boundaries are blurred or contrast with the terrain is low. Traditional superpixel segmentation methods often perform poorly in these weak boundary areas, leading to superpixels that cross over both cloud and non-cloud areas, thus reducing the accuracy and efficiency of subsequent processing. To address this issue, this study proposes a method of superpixel segmentation for automatic cloud boundary detection in satellite imagery based

on local adaptive distance. This method specifically targets the challenge of identifying cloud boundaries in satellite imagery by dynamically adjusting segmentation strategies to adapt to subtle changes between clouds and the terrain, optimizing the generation of superpixels to ensure they strictly conform to cloud boundaries, significantly reducing semantic errors due to under-segmentation of weak boundaries.

## 2.1 Definition of local adaptive distance metric

In the context of automatic cloud detection and removal in satellite imagery, precise segmentation of cloud boundaries is one of the core tasks, especially in areas where contrast between clouds and the ground is low. Traditional superpixel algorithms, which use a globally consistent distance metric, often fail to effectively identify subtle boundaries in low contrast areas. The method proposed in this study for superpixel segmentation sensitive to cloud boundaries based on local adaptive distance emphasizes sensitivity and discriminative ability for weak boundaries. This method adaptively amplifies the feature differences that are less contrasted between cloud and non-cloud areas, thereby enhancing detection capability for weak boundaries. Figure 1 shows the advantage of local adaptive distance measurement over non-local distance measurement.

In the definition of the local adaptive distance metric, the representation space of the image  $U$  is defined as a multi-dimensional feature space, which includes original image features such as color and brightness, as well as other features that may be extracted based on satellite image characteristics (e.g., infrared and multispectral data). Given the digital image  $U$ , with a total number of pixels  $V$ , the result of superpixel segmentation is denoted as  $A=\{A_j\}_{j=1}^J$ , where  $J$  is the predetermined number of superpixels. The feature space of image  $U$  is defined as  $U=\{\{Z_k\}_{k=1}^n, \{Z_a, Z_b\}\}$ , where  $n$  is the dimension of the feature space,  $Z_k$  represents the  $k$ -th feature channel, while  $Z_a$  and  $Z_b$  represent spatial location coordinates. Each pixel  $u$  in this space is represented as a vector  $o_u=[z_{1,u}, \dots, z_{n,u}, z_{a,u}, z_{b,u}]$ .

Considering the feature differences and spatial connectivity between cloud and non-cloud areas, to quantify the distance more precisely between pixels and superpixels, this study proposes the following distance function. This distance function combines feature dissimilarity and spatial distance. The function introduces a balancing factor to adjust the weight between the feature homogeneity term and the regularity constraint term to adapt to the characteristics of different areas. Particularly in weak contrast areas at cloud boundaries, by increasing the weight of feature dissimilarity in the distance measurement, more effective identification and maintenance of these subtle boundaries can be achieved, while ensuring the spatial connectivity and shape regularity of superpixels. Specifically, the feature homogeneity term is represented by  $F_z(o_u, A_j)$ , the regularity constraint term by  $F_t(o_u, A_j)$ , the number of feature dimensions of the input image by  $n$ , and the balancing factor between the feature homogeneity term and the regularity constraint term by  $\eta$ . Given pixel  $o_u$  and superpixel  $A_j$ , their distance can be calculated as follows:

$$F_{MX}(o_u, A_j) = \sqrt{F_z^2(o_u, A_j) + \eta \cdot n \cdot F_t^2(o_u, A_j)} \quad (1)$$

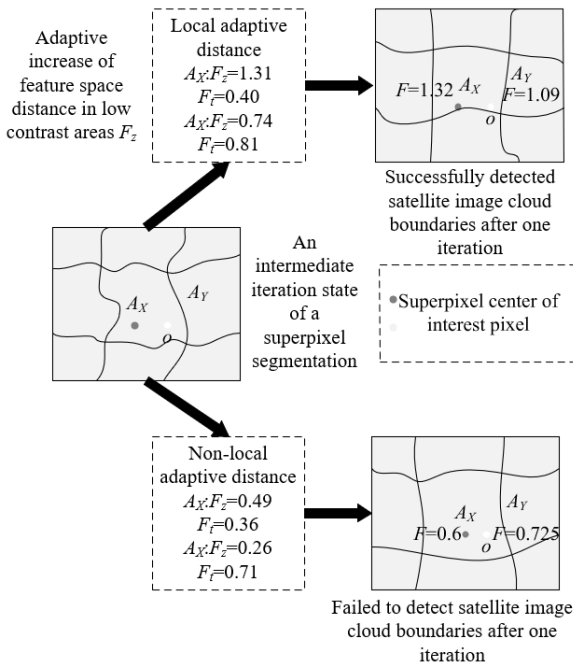
In the method of superpixel segmentation for automatic cloud boundary detection in satellite imagery, the feature homogeneity term is crucial for ensuring the consistency of pixels within a superpixel and their precise alignment with actual cloud boundaries. The proposed method uses a feature homogeneity term based on local adaptive distance, emphasizing the consistency of features such as grayscale, color, or infrared data within superpixels and ensuring that the boundaries of the superpixels tightly adhere to the natural edges of cloud layers. From a mathematical perspective, by performing Z-score normalization of the image within local regions, the contrast in low-contrast areas can be effectively increased, easing the difficulty of boundary segmentation. This normalization not only ensures the uniformity of features but also eliminates the impact of different feature scales on distance calculations, allowing the algorithm to handle different feature representations more flexibly in satellite imagery, thus enhancing the accuracy and efficiency of cloud boundary detection. Assuming that the mean and standard deviation of all pixels in superpixel  $A_j$  on feature channel  $Z_k$  are represented by  $\omega_{k,j}$  and  $\delta_{k,j}$  respectively,  $\omega_{k,j}$  and  $\delta_{k,j}$  are defined as  $\omega_{k,j} = |A_j|^{-1} \sum_{u \in A_j} z_{k,j}$ ,  $\delta_{k,j} = \sqrt{|A_j|^{-1} \sum_{u \in A_j} (z_{k,j} - \omega_{k,j})^2}$ . The following equation defines the feature homogeneity term:

$$F_z(o_u, A_j) = \sqrt{\sum_{k=1}^n q_{k,j}^2 A_j \cdot f_k^2(o_u, A_j)} \quad (2)$$

$$f_k(o_u, A_j) = |z_{k,u} - \omega_{k,j}| \quad (3)$$

$$w_{j,k}(X_k) = (\sigma_{j,k})^{-1} \quad (4)$$

In the process of superpixel segmentation sensitive to automatic cloud boundaries in satellite imagery, applying a global constraint to the feature homogeneity term is a key step in ensuring segmentation quality. The purpose of this constraint is to reduce the jittering of superpixel boundaries in flat areas caused by minor feature perturbations, which is particularly noticeable when the feature standard deviation approaches zero, and has no substantive meaning for actual



**Figure 1.** Comparison of local adaptive distance measurement to non-local distance measurement

image segmentation. Therefore, this study adopts an adaptive global constraint method by setting a minimum limit on the feature standard deviation of superpixels, avoiding unstable boundaries caused by overly sensitive feature weights. Specifically, this global constraint uses the cube root of the third moment of the feature standard deviation of superpixels as a threshold. This threshold not only considers the complexity of the image content but also achieves a balance in the sensitivity to features in different areas of the image through the control parameter  $l$ . Assuming the initial standard deviation of features within the superpixel division is represented by  $\delta_0$ , the calculation formula is:

$$Sg_k = l \cdot \sqrt[3]{\frac{1}{J} \sum_{j=1}^J ((\delta_0)_{k,j})^3} \quad (5)$$

After imposing a minimum threshold limit on the standard deviation of each feature channel of the satellite image superpixels, the local adaptive weights can be set as follows:

$$q_{k,j}(A_j) = (\text{MAX}(\delta_{k,j}, Sg_k))^{-1} \quad (6)$$

Assuming that the mean spatial coordinates  $a$  and  $b$  of all pixels within superpixel  $A_j$  are represented by  $\omega_{a,j}$  and  $\omega_{b,j}$ , and the normalization factor by  $X$ . Further, set the spatial distance metric  $F_i$ :

$$F_i(o_u, A_j) = \sqrt{\frac{(z_{a,u} - \omega_{a,j})^2 + (z_{b,u} - \omega_{b,j})^2}{X^2}} \quad (7)$$

The final distance consists of two parts,  $F_i$  and  $F_z$ , balanced by  $\eta_0$ . Since  $F_z^2$  increases with the increase in feature dimensions,  $\eta_0$  is set as the product of the feature dimension  $n$  and the adjustable parameter  $\eta$ . The final distance between pixel  $o_u$  and superpixel  $A_j$  can be defined as follows:

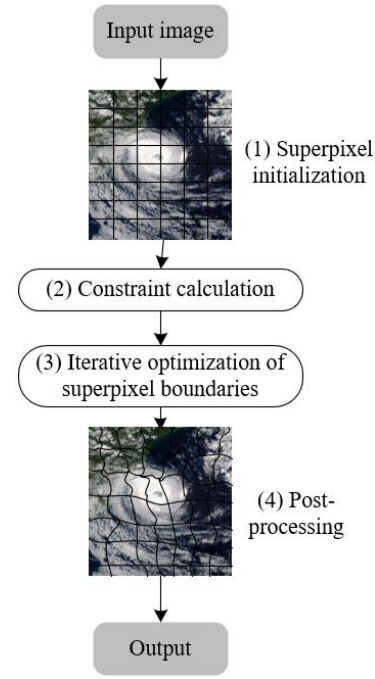
$$D_{LA}^2(p_i, X_k) = \sum_{j=1}^v \frac{(c_{j,i} - \mu_{j,k})^2}{(\max(\sigma_{j,k}, Th_j))^2} + \lambda \cdot v \cdot \frac{(c_{x,i} - \mu_{x,k})^2 + (c_{y,i} - \mu_{y,k})^2}{A^2} \quad (8)$$

## 2.2 Automatic cloud boundary superpixel generation in satellite imagery

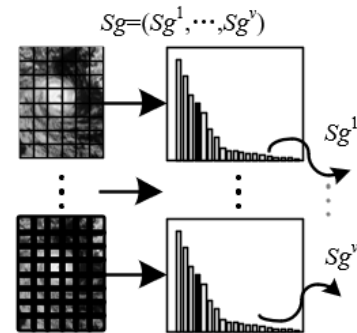
To improve the feature homogeneity of superpixels, this paper proposes a satellite imagery automatic cloud boundary superpixel segmentation model based on morphological contour evolution. Below, the basic principles of the segmentation model are detailed. Figure 2 shows the process flow for the generation of automatic cloud boundary superpixels in satellite imagery.

The model initializes using a regular grid method, similar to the USEAQ and SCAC algorithms, where the initial side length of the superpixels is set to  $T=VN/J$ , calculated based on the total number of pixels  $V$  and the predetermined number of superpixels  $J$ . Further, global constraint calculations are carried out, which involve setting initial rounds of superpixel segmentation and then performing a global analysis of features and spatial constraints across the entire image to ensure that

the superpixels are not only homogeneous in features but also regular in shape and size. By calculating the standard deviation of features within each superpixel region and applying the previously mentioned minimum value constraints, segmentation errors caused by overly sensitive features can be effectively avoided. Additionally, the global constraints consider the spatial relationships between superpixels, using normalized spatial distances to maintain the shape regularity of superpixels. Figure 3 illustrates the constraint calculation process.



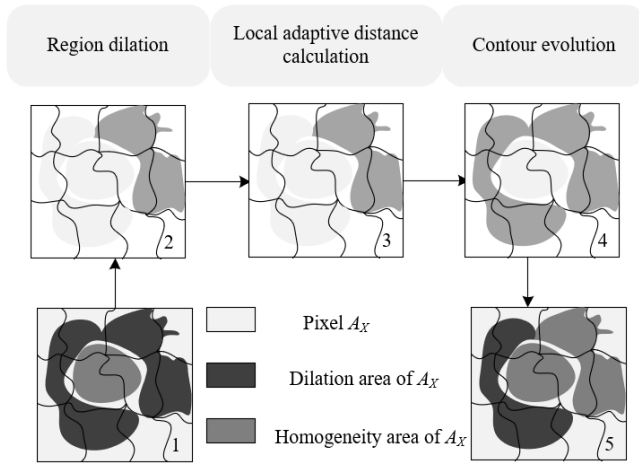
**Figure 2.** Process flow for automatic cloud boundary superpixel generation in satellite imagery



**Figure 3.** Constraint calculation process

The iterative superpixel generation step is the core component of the model, crucial for effectively expanding and optimizing the shape of superpixels to map and distinguish between cloud and non-cloud areas more accurately. This step starts from the initialized regular superpixel grid, using binary morphological dilation operations to define the evolution search area for each superpixel. Morphological dilation extends the current boundaries of the superpixels using a defined structural element, achieving uniform expansion of the boundaries, which not only maintains balanced distances but also adapts to irregular and elongated boundary shapes. The evolution operation for each superpixel is confined within a predefined maximum range, typically several multiples of the

initial side length ( $M \times T$ ), which ensures the controllability of operations and avoids the inaccuracies brought by excessive dilation. Figure 4 demonstrates the iterative optimization process for superpixel cloud boundary generation.



**Figure 4.** Iterative optimization process for superpixel cloud boundary generation

In each iteration, by comparing the local adaptive distance between pixels within neighboring areas and the current superpixel, pixels that are feature-similar and closer in distance are selectively incorporated into the superpixel. This method is particularly suitable for handling cloud boundary-sensitive areas in satellite imagery, as it allows superpixels to adaptively respond to subtle changes between clouds and non-clouds, thereby improving the accuracy of cloud boundary detection. Experimental results show that selecting an appropriate  $M$  value can strike a good balance between segmentation precision and algorithm efficiency, effectively supporting the fine-tuning needs of superpixels in cloud detection and removal tasks. This iterative generation process not only enhances the feature homogeneity of the superpixels but also optimizes the shape and functionality of the superpixels through precise control of the evolution process, ensuring the efficiency and accuracy of satellite imagery segmentation.

Finally, post-processing is conducted, mainly aimed at enhancing the connectivity and consistency of the superpixel segmentation results, especially addressing connectivity issues that may arise during the processing of larger-sized superpixels. Through post-processing, the model focuses on correcting those connected areas that are too small and may have become isolated due to improper segmentation. Specific operations include merging these small areas with larger superpixels that are nearest in the feature space, thus ensuring greater homogeneity of features within each superpixel and maintaining the spatial connectivity of the superpixels.

### 3. AUTOMATIC CLOUD REMOVAL IN SATELLITE IMAGERY BASED ON GAN

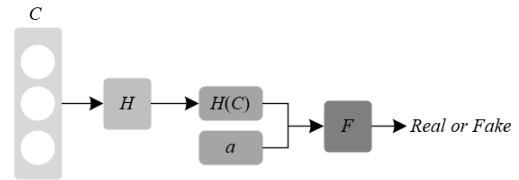
Figure 5 shows the traditional schematic diagram of a GAN structure. Further, this paper proposes an automatic cloud removal model for satellite imagery based on GAN. The entire network framework is designed to include three key sub-networks: the cloud removal image generation network  $H_i$ , the transmission mapping network  $H_s$ , and the exposure

estimation network  $R$ . The cloud removal image generation network  $H_i$  is responsible for receiving the style-transformed synthetic image output by the style conversion network  $H_i$ , which simulates the presence of clouds in satellite imagery, and outputs the image  $U_i$  after cloud removal through network processing. The task of the transmission mapping network  $H_s$  is to estimate the transmission map  $U_t$  from cloud-covered images to cloud-free images, reflecting the transmission characteristics between cloud-covered and non-cloud areas in the image. The exposure estimation network  $R$  is used to assess the differences in exposure between cloud and non-cloud areas, simulating the occlusion effects caused by clouds. This GAN-based model, through these three collaboratively working networks, can not only effectively remove clouds from satellite imagery but also restore the surface details of the areas covered by clouds, greatly improving the quality and utility of the images after cloud removal, supporting more accurate ground observation and analysis. The following equation gives the automatic cloud area expression in satellite imagery:

$$U_t = U_i * U_z + (1 - U_i) * U_r \quad (9)$$

After training all the generator networks, the trained automatic cloud removal images for satellite imagery are obtained, and the following formula is used to construct the automatic cloud removal image for satellite imagery:

$$U_{out} = \frac{U_i - U_r(1 - U_i)}{U_i} \quad (10)$$



**Figure 5.** Schematic diagram of GAN structure

In the automatic cloud removal model for satellite imagery based on GAN, the main objective of the cloud-free image generation network is to generate pseudo-real images without cloud coverage. These images need to accurately reflect the distribution of illumination on the ground, while maintaining the style and color attributes of objects within the scene. To ensure the stability of the  $H_i$  network and prevent overfitting, this study adopts the FFA-Net architecture, which effectively preserves the edges, contours, and geometric structures of images using channel and pixel attention mechanisms. Compared to other style transformation networks, FFA-Net avoids overfitting more effectively during training, thus producing more ideal image effects. This method is particularly suitable for the task of cloud removal in satellite imagery, as it needs to accurately restore the information beneath the clouds while maintaining the overall consistency and realism of the image. For  $H_i$ , the  $L1$  norm is used as the loss function:

$$M_{L1}(H_i) = R_{z \sim O_{DA}(z), I \sim O_{DA}(I)} \left[ \|H_i(z) - i\|_1 \right] \quad (11)$$

In the model, the transmission mapping generator  $H_s$  is a crucial component whose primary function is to generate a

transmission map from the input cloud-covered satellite imagery. This map illustrates the transition from cloud-covered to cloud-removed states.  $H_s$  adopts an encoder-decoder end-to-end network structure and is trained in a supervised learning manner, aiming to produce a transmission map closely resembling the true cloud-removed image. The network is based on the CycleGAN architecture with the addition of 10 residual blocks to enhance learning capability and network stability. During training, each input of a cloud-covered image, through the collaborative work of  $H_i$  and  $H_s$ , generates the cloud-removed image  $L_v$  and its corresponding transmission map. The loss function of  $H_s$  mainly consists of an  $L1$  loss, which calculates pixel consistency, while the edge loss is used to reinforce the continuity and smooth transition of image edges and contours, ensuring that the generated transmission map is visually more natural and the edges are smoother, helping to improve the accuracy and visual quality of cloud removal. The process can be represented as:

$$L_v = H_i(H_i(T_v)) \quad (12)$$

Assuming the Sobel operator is represented by  $R$  and the discriminator trained using adversarial loss is represented by  $F_l$ , the loss function of  $H_s$  is expressed as:

$$M_{Ll}(H_l) = R_{i \sim O_{Dl}(l), l \sim O_{Dl}(L)}[\|H_l(i) - l\|_1] \quad (13)$$

$$M_{ED}(H_l) = R_{i \sim O_{Dl}(l), l \sim O_{Dl}(L)}[\|\gamma H_l(i) - \gamma l\|_1] \quad (14)$$

$$M(H_l, F_l) = \alpha_1 M_{Ll}(H_l) + \alpha_2 M_{ED}(H_l) + M_{AD}(H_l, F_l, L) \quad (15)$$

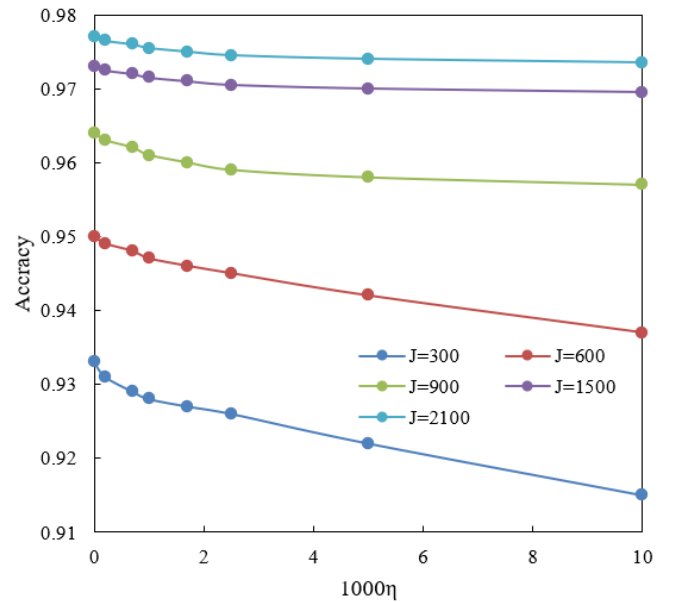
In the model, the exposure estimation network  $R$  plays a critical role, mainly functioning to estimate the lighting conditions under cloud-covered areas. This network takes the input cloud-covered satellite imagery and outputs a predicted three-channel light map  $r = \{r_e, r_h, r_y\}$  by analyzing each channel in the RGB space, similar to traditional overexposure image estimation methods. Unlike conventional full-image processing methods, to be compatible with the transmission mapping generator and considering processing efficiency and flexibility, the  $R$  network is trained and applied on  $32 \times 32$  image blocks. This design enables  $R$  to effectively estimate lighting across different resolution image blocks, beneficial for the overall flexibility and accuracy of the model.

Inspired by the U-Net architecture,  $R$ 's network design includes four downsampling and upsampling blocks to accommodate the processing needs of small image blocks. The convolution blocks used within the network have their kernel sizes and strides adjusted to  $2 \times 2$  and 2, respectively, to suit small patch sizes, and the original pooling layers of U-Net have been removed. As an estimation regression type of network,  $R$  structurally adopts two fully connected layers in place of traditional blocks. Additionally,  $R$  is trained using an  $L1$  loss function to optimize its estimation accuracy of lighting conditions under cloud effects in satellite imagery. The expression for this is:

$$M_{Ll}(R) = \frac{1}{V} \left[ \sum_{u=1}^V \|\{r'_e, r'_h, r'_y\} - \{r_e, r_h, r_y\}\| \right] \quad (16)$$

#### 4. EXPERIMENTAL RESULTS AND ANALYSIS

According to the data shown in Figure 6, we can observe the accuracy performance of the automatic cloud boundary sensitive superpixel segmentation in satellite imagery under different settings of the number of superpixels  $J$  and the balance factor  $\eta$  for feature homogeneity and regularity constraints. The data shows that when  $\eta$  is 0, meaning the balance factor of feature homogeneity and regularity constraints is not considered, the accuracy increases from 0.933 to 0.977 as the number of superpixels  $J$  increases from 300 to 2100, indicating that increasing the number of superpixels enhances the accuracy of superpixel segmentation. However, as the value of  $\eta$  increases from 0 to 10, there is a general trend of slight decline in segmentation accuracy. For example, at  $J=1500$ , when  $\eta$  increases from 0 to 10, the accuracy decreases from 0.973 to 0.9695. This trend is evident across all superpixel number settings, suggesting that a higher value of  $\eta$  might have a slightly negative impact on segmentation accuracy during the superpixel segmentation process. Analysis indicates that the data in the figure reflects that appropriately increasing the number of superpixels can significantly improve segmentation accuracy in the application of automatic cloud boundary sensitive superpixel segmentation, mainly because more superpixels help to delineate the complex features of cloud boundaries more finely. Additionally, although increasing the value of  $\eta$  is intended to enhance the precision of segmentation by strengthening the constraints on feature homogeneity and regularity, experimental data shows that too high a value of  $\eta$  might lead to a slight decrease in segmentation performance due to over-smoothing of boundaries.

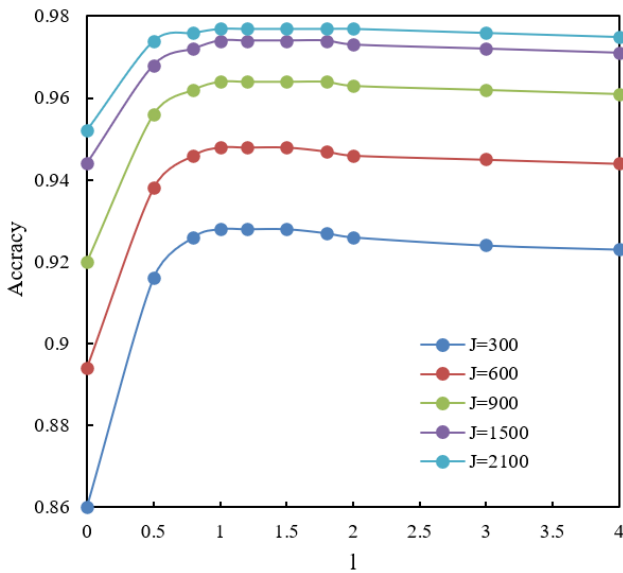


**Figure 6.** The impact of the value of parameter  $\eta$  on the accuracy of automatic cloud boundary sensitive superpixel segmentation in satellite imagery

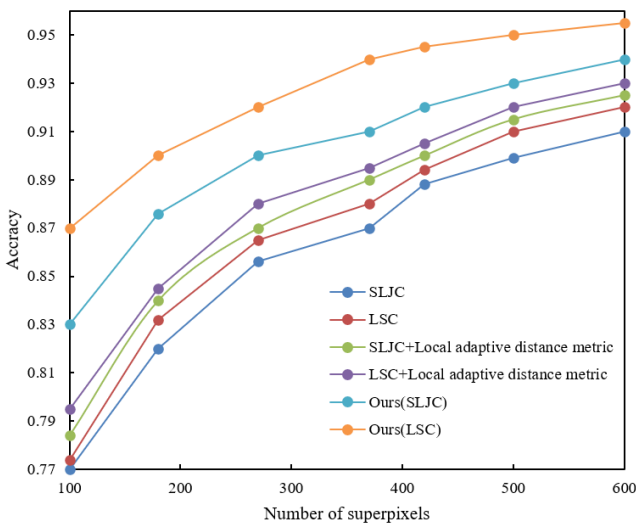
According to the data in Figure 7, we can observe the impact of the control parameter  $l$  on the accuracy of automatic cloud boundary sensitive superpixel segmentation in satellite imagery across different numbers of superpixels  $J$ . From  $l = 0$  to  $l = 1$ , there is a significant improvement in segmentation accuracy across all superpixel settings. For example, at  $J = 1500$ , accuracy increases from 0.944 to 0.974, showing that as



the value of  $l$  increases, there is a notable improvement in segmentation performance. However, when the value of  $l$  exceeds 1, the increase in accuracy becomes more gradual, and at  $J=2100$ , there is a slight decline from 0.977 at  $l=1$  to 0.975 at  $l=4$ , indicating that continuing to increase  $l$  beyond a certain threshold might have a minor negative impact or no significant improvement on segmentation performance. These results suggest that appropriately increasing the control parameter  $l$  can significantly enhance the accuracy of superpixel segmentation of cloud boundaries in satellite imagery, especially under medium to high superpixel number settings. This improvement is mainly due to the increase in  $l$  value, making superpixel segmentation more focused on the adaptive features of local areas, effectively enhancing the recognition accuracy of cloud boundary regions.



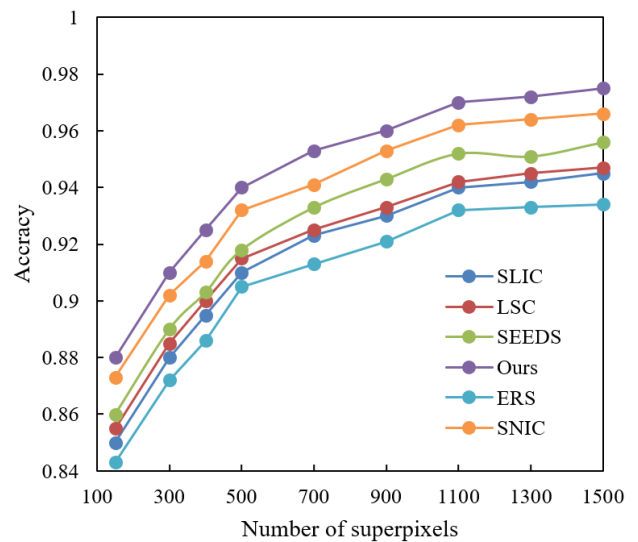
**Figure 7.** The impact of the value of parameter  $l$  on the accuracy of automatic cloud boundary sensitive superpixel segmentation in satellite imagery



**Figure 8.** Impact of different methods on the accuracy of automatic cloud boundary sensitive superpixel segmentation in satellite imagery

According to the data shown in Figure 8, we can analyze and compare the performance of different superpixel segmentation methods on the accuracy of automatic cloud

boundary sensitive superpixel segmentation in satellite imagery. From the figure, it is evident that the accuracy of baseline models SLJC and LSC improves with an increase in the number of superpixels, indicating that more detailed superpixel segmentation helps better recognize cloud boundaries. However, the accuracy of both the SLJC and LSC models significantly increases after introducing the local adaptive distance measure. For example, at a superpixel count of 500, SLJC increased from 0.899 to 0.915, and LSC from 0.91 to 0.92. Moreover, the methods proposed in this study, "Ours (SLJC)" and "Ours (LSC)," display higher accuracy across all superpixel counts, particularly at 600 superpixels, where Ours (SLJC) and Ours (LSC) reached accuracies of 0.94 and 0.955 respectively. These results clearly demonstrate the effectiveness of the superpixel segmentation technique based on local adaptive distance in improving the accuracy of cloud boundary recognition in satellite imagery. Compared to traditional SLJC and LSC methods, the introduction of local adaptive distance measurement significantly enhances segmentation accuracy, mainly due to the technique's ability to handle complex local features and variations more precisely in the imagery, thus better adapting to the different characteristics of cloud and non-cloud areas. Ultimately, the improved models proposed in this study further enhance segmentation results, showing the potential and practicality of this technology, providing an effective tool for future satellite image processing.



**Figure 9.** Impact of different methods on the accuracy of automatic cloud boundary sensitive superpixel segmentation in satellite imagery

Figure 9 displays the performance of various superpixel segmentation algorithms in the task of automatic cloud boundary sensitive superpixel segmentation in satellite imagery. From the figure, it can be observed that the segmentation accuracy of all algorithms generally increases with the number of superpixels. Among these, our method shows higher accuracy across all superpixel settings, gradually increasing from the lowest at 0.88 with 150 superpixels to 0.975 with 1500 superpixels. In comparison, other traditional algorithms like SLIC, LSC, SEEDS, ERS, and SNIC, although also improving in accuracy with an increase in superpixel numbers, perform less effectively at equivalent superpixel counts than our method. For example, at 1500 superpixels, our method's accuracy is 0.975, while the closest algorithm, SNIC,

has an accuracy of 0.966, and other methods like SLIC and LSC are even lower. These results clearly indicate the significant advantages of the superpixel segmentation technique based on local adaptive distance measurement developed in this study for the task of automatic cloud boundary sensitive segmentation in satellite imagery. By meticulously analyzing the complex relationships between pixels and adaptively determining cloud boundaries, our method not only enhances the precision of cloud detection but also achieves more accurate boundary delineation at higher superpixel counts.

**Table 1.** RMSE Results for automatic cloud removal in satellite imagery using different methods

<i>Methods</i>		<b>Training Set</b>	<b>Test Set</b>
<i>K-Means Clustering</i>	Light Cloud Coverage Images	12.362	15.234
	Moderate Cloud Coverage Images	6.895	9.562
	Heavy Cloud Coverage Images	8.562	11.254
<i>FCN</i>	Light Cloud Coverage Images	15.241	17.895
	Moderate Cloud Coverage Images	6.124	9.562
	Heavy Cloud Coverage Images	7.652	10.235
<i>Kernel PCA</i>	Light Cloud Coverage Images	10.234	12.321
	Moderate Cloud Coverage Images	7.215	10.248
	Heavy Cloud Coverage Images	7.854	10.258
<i>One-Class SVM</i>	Light Cloud Coverage Images	6.592	9.568
	Moderate Cloud Coverage Images	3.785	6.658
	Heavy Cloud Coverage Images	4.125	7.581
<i>Deep Forest</i>	Light Cloud Coverage Images	7.889	10.234
	Moderate Cloud Coverage Images	3.142	6.358
	Heavy Cloud Coverage Images	3.899	6.598
<i>Ours</i>	Light Cloud Coverage Images	6.012	9.458
	Moderate Cloud Coverage Images	3.124	6.654
	Heavy Cloud Coverage Images	3.879	6.623

Table 1 displays the performance of various methods in the task of automatic cloud removal from satellite imagery, using Root Mean Square Error (RMSE) as the evaluation criterion. The data show that the effectiveness of each method varies under different cloud cover conditions (light, moderate, heavy). Specifically, One-Class SVM and Deep Forest exhibit excellent performance across all types of cloud cover, particularly in moderate and heavy cloud-covered images, where the RMSE values are relatively low, indicating significant effectiveness in handling images with medium to heavy cloud coverage. In contrast, FCN and K-means Clustering show higher RMSE in light cloud coverage images, indicating less than ideal performance. Particularly, the FCN method has the highest test set RMSE reaching 17.895 in light

cloud coverage images, possibly due to these methods being insufficiently sensitive in handling low cloud cover densities. The proposed method demonstrates stable performance across all categories of cloud coverage, especially in heavy cloud coverage images, where its RMSE results are close to those of One-Class SVM and Deep Forest, showing good performance. These results underscore the effectiveness and practicality of the automatic cloud removal method for satellite imagery based on GAN proposed in this study. By integrating local adaptive distance measurement within the GAN framework, this approach is not only theoretically innovative but also demonstrates excellent practical outcomes in restoring surface information in cloud-covered areas, particularly achieving lower errors in handling heavy cloud-covered images, ensuring the naturalness and accuracy of the cloud removal results.

**Table 2.** Performance metrics describing super-resolution performance at a magnification factor of  $\times 2$  on the test set

<i>Methods</i>	<b>Light Cloud Coverage Images</b>		<b>Moderate Cloud Coverage Images</b>		<b>Heavy Cloud Coverage Images</b>		<b>All</b>	
	<i>PSNR/SSIM</i>	<i>PSNR/SSIM</i>	<i>PSNR/SSIM</i>	<i>PSNR/SSIM</i>	<i>PSNR/SSIM</i>	<i>PSNR/SSIM</i>	<i>PSNR/SSIM</i>	<i>PSNR/SSIM</i>
<i>K-Means Clustering</i>	19.265	0.7785	20.154	0.7895	20.125	0.7894	20.562	0.7985
<i>FCN</i>	19.985	0.7895	21.369	0.7954	21.125	0.7959	21.124	0.7963
<i>Kernel PCA</i>	20.124	0.7985	21.598	0.7985	21.365	0.8152	21.895	0.8124
<i>One-Class SVM</i>	20.568	0.7962	21.895	0.7845	22.398	0.7985	21.963	0.8136
<i>Deep Forest</i>	20.689	0.8125	22.314	0.8156	22.569	0.8236	22.658	0.8265
<i>Ours: 24.54128235</i>								

Table 2 provides the performance metrics for image super-resolution processing on the test set using different methods, specifically for a magnification factor of  $\times 2$ , covering Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) across satellite images with light, moderate, and heavy cloud coverage. The table shows that the performance of each method varies with the degree of cloud coverage, but generally, both PSNR and SSIM are slightly higher in heavily cloud-covered images compared to light and moderate cloud coverage. Particularly, the Deep Forest method performs excellently under all cloud coverage conditions, with PSNR at 22.569 and SSIM at 0.8236 for heavy cloud coverage images. However, the method proposed in this study shows a significant improvement in super-resolution performance, with an overall PSNR value of 24.541. This result indicates a significant advantage in terms of image quality improvement and detail restoration. The analysis clearly demonstrates that the super-resolution method developed based on GAN exhibits outstanding performance in satellite image processing. By integrating local adaptive distance measurement, this method not only enhances the visual quality of images but also achieves higher standards in detail restoration, particularly showing exceptional capabilities in image magnification and clarity enhancement.

Table 3 details the performance of different methods for automatic cloud removal from satellite imagery at a magnification factor of  $\times 2$ , using RMSE as the metric. The table shows variations in performance across different degrees of cloud coverage, but overall, Deep Forest and One-Class SVM perform notably well, especially in moderate cloud

coverage images where the RMSE values are lower. Notably, the proposed method exhibits exceptional removal effects across all cloud coverage conditions, particularly in moderate cloud coverage images with an RMSE of only 12.8987, significantly lower than other methods, demonstrating its effective cloud removal capabilities. Moreover, in light and heavy cloud coverage images, the RMSE values for our method are 26.4256 and 16.3256 respectively, also showing superior performance compared to other methods. These data results emphasize the effectiveness of the satellite imagery automatic cloud removal algorithm based on GAN technology. By integrating an advanced GAN framework and local adaptive distance measurement, our method not only excels in handling moderate cloud coverage images but also maintains low error rates under light and heavy cloud coverage conditions, showing good adaptability and robustness.

**Table 3.** RMSE metrics describing automatic cloud removal performance at a magnification factor of  $\times 2$  on the test set

Methods	Light Cloud Coverage Images	Moderate Cloud Coverage Images	Heavy Cloud Coverage Images	All
<i>K-Means Clustering</i>	52.132	51.2346	51.2365	51.2351
	25.368	22.3255	22.3658	21.235
	31.248	25.6986	25.6984	25.314
<i>FCN</i>	57.263	52.3142	52.1341	52.314
	22.361	18.6948	18.9642	18.6586
	26.598	22.3143	22.3588	22.1354
<i>Kernel PCA</i>	42.315	36.5681	37.1549	36.5987
	23.568	18.3258	18.6253	18.6257
	27.895	26.3289	24.2658	25.3159
<i>One-Class SVM</i>	39.258	34.2653	35.2686	36.1248
	22.314	17.2651	18.1458	16.5689
	25.362	21.2368	22.1351	22.3145
<i>Deep Forest</i>	41.258	36.2659	36.9851	36.9851
	21.268	16.9859	16.5695	16.5984
	25.369	22.3151	22.1458	22.1358
	<i>Ours: 26.4256/12.8987/16.3256</i>			

## 5. CONCLUSION

This paper's research primarily revolves around two core technologies: superpixel segmentation of satellite imagery sensitive to cloud boundaries based on local adaptive distance and an automatic cloud removal algorithm for satellite imagery based on GAN. By deeply analyzing complex relationships between pixels, the first technology adaptively determines cloud boundaries, enhancing cloud detection accuracy; while the second part utilizes the generative capabilities of GAN to effectively restore surface information in cloud-covered areas, enhancing the application value of satellite imagery.

Experimental results demonstrate that appropriate adjustments of the balance factor  $\eta$  for feature homogeneity and regularity constraints, as well as the control parameter  $l$ , significantly enhance the accuracy of superpixel segmentation. Moreover, compared to traditional methods, the technologies proposed in this paper show superior performance in both superpixel segmentation precision and cloud removal efficiency. Particularly in automatic cloud removal, whether in RMSE or super-resolution performance metrics, our method significantly outperforms other comparative methods, especially effective in handling images with moderate to heavy cloud coverage.

Overall, this research not only provides innovative theoretical insights but also demonstrates significant practical benefits, particularly in geographic information systems and environmental monitoring fields, greatly enhancing the usability and accuracy of satellite data. However, the study's limitations mainly manifest in the potential over-reliance on experimental settings for parameter selection, which in different real-world application scenarios might require further adjustment and validation.

Future research directions could include further optimizing the parameter self-adjustment mechanisms of the algorithm to improve its generalizability across various environments. Additionally, considering the diversity and dynamic changes in cloud coverage, future studies could explore integrating time-series data for cloud detection and removal techniques to adapt to more complex and variable real-world application needs. Ultimately, these studies will further advance remote sensing technology, providing stronger technical support for precise Earth observation.

## REFERENCES

- [1] Sasaki, K., Sekine, T., Emery, W. (2024). Enhancing the detection of coastal marine debris in very high resolution satellite imagery via unsupervised domain adaptation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 6014-6028. <https://doi.org/10.1109/JSTARS.2024.3364165>
- [2] Camarretta, N., Pearse, G.D., Steer, B.S.C., McLay, E., Fraser, S., Watt, M.S. (2024). Automatic detection of phytophthora pluvialis outbreaks in radiata pine plantations using multi-scene, multi-temporal satellite imagery. *Remote Sensing*, 16(2): 338. <https://doi.org/10.3390/rs16020338>
- [3] Wang, Z. (2023). Flash flood management topographic survey assisted by satellite imagery mapping technology. In *Proceedings of the 2023 International Conference on Intelligent Sensing and Industrial Automation*, New York, NY, United States, pp. 1-5. <https://doi.org/10.1145/3632314.3632338>
- [4] Ma, J., Shen, H., Cai, Y., Zhang, T., Su, J., Chen, W.H., Li, J. (2023). UCTNet with dual-flow architecture: snow coverage mapping with Sentinel-2 satellite imagery. *Remote Sensing*, 15(17): 4213. <https://doi.org/10.3390/rs15174213>
- [5] Pitonak, R., Mucha, J., Dobis, L., Javorka, M., Marusin, M. (2022). Cloudsatnet-1: FPGA-based hardware-accelerated quantized CNN for satellite on-board cloud coverage classification. *Remote Sensing*, 14(13): 3180. <https://doi.org/10.3390/rs14133180>
- [6] Moschos, E., Stegner, A., Schwander, O., Gallinari, P. (2020). Classification of eddy sea surface temperature signatures under cloud coverage. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 3437-3447. <https://doi.org/10.1109/JSTARS.2020.3001830>
- [7] Xu, X., Zeng, Y., Yu, X., Liu, G., Yue, Z., Dai, J., Zhu, Y. (2022). Identification of supercooled cloud water by FY-4A satellite and validation by CALIPSO and airborne detection. *Remote Sensing*, 15(1): 126. <https://doi.org/10.3390/rs15010126>
- [8] Gómez, J.E.M., Cheng, A.M. (2022). Work-in-progress: Real-time on-board processing for cloud detection in



- FACSAT-2 multispectral satellite imagery. In 2022 IEEE Real-Time Systems Symposium (RTSS), Houston, TX, USA, pp. 499-502. <https://doi.org/10.1109/RTSS55097.2022.00051>
- [9] Haynes, J. M., Noh, Y.J., Miller, S.D., Haynes, K.D., Ebert-Uphoff, I., Heidinger, A. (2022). Low cloud detection in multilayer scenes using satellite imagery with machine learning methods. *Journal of Atmospheric and Oceanic Technology*, 39(3): 319-334. <https://doi.org/10.1175/JTECH-D-21-0084.1>
- [10] Gupta, R., Nanda, S.J. (2020). Improved framework of many-objective evolutionary algorithm to handle cloud detection problem in satellite imagery. *IET Image Processing*, 14(17): 4795-4807. <https://doi.org/10.1049/iet-ipr.2020.0535>
- [11] Zhang, Z., Zheng, H., Cao, J., Feng, X., Xie, G. (2022). FRS-Net: An efficient ship detection network for thin-cloud and FOG-covered high-resolution optical satellite imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15: 2326-2340. <https://doi.org/10.1109/JSTARS.2022.3227322>
- [12] Li, Z., Shen, H., Weng, Q., Zhang, Y., Dou, P., Zhang, L. (2022). Cloud and cloud shadow detection for optical satellite imagery: Features, algorithms, validation, and prospects. *ISPRS Journal of Photogrammetry and Remote Sensing*, 188: 89-108. <https://doi.org/10.1016/j.isprsjprs.2022.03.020>
- [13] Zhai, H., Xue, L. (2024). AMCD-Net: An effective attention aided multilevel cloud detection network for optical satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1-12. <https://doi.org/10.1109/TGRS.2024.3372589>
- [14] Parmes, E., Rauste, Y., Molinier, M., Andersson, K., Seitsonen, L. (2017). Automatic cloud and shadow detection in optical satellite imagery without using thermal bands—Application to Suomi NPP VIIRS images over Fennoscandia. *Remote Sensing*, 9(8): 806. <https://doi.org/10.3390/rs9080806>
- [15] Molnár, T., Király, G. (2024). Forest disturbance monitoring using cloud-based sentinel-2 satellite imagery and machine learning. *Journal of Imaging*, 10(1): 14. <https://doi.org/10.3390/jimaging10010014>
- [16] Hu, C., Zhang, Z., Tang, P. (2023). Research on multispectral satellite image cloud and cloud shadow detection algorithm of domestic satellite. *National Remote Sensing Bulletin*, 27(3): 623-634. <https://doi.org/10.11834/jrs.20211209>
- [17] Pang, S., Sun, L., Tian, Y., Ma, Y., Wei, J. (2023). Convolutional neural network-driven improvements in global cloud detection for Landsat 8 and transfer learning on sentinel-2 imagery. *Remote Sensing*, 15(6): 1706. <https://doi.org/10.3390/rs15061706>
- [18] Yin, Z., Ling, F., Foody, G. M., Li, X., Du, Y. (2020). Cloud detection in Landsat-8 imagery in Google Earth Engine based on a deep convolutional neural network. *Remote Sensing Letters*, 11(12): 1181-1190. <https://doi.org/10.1080/2150704X.2020.1833096>
- [19] Tao, R., Zhang, Y., Wang, L., Liu, Q., Wang, J. (2021). U-High resolution network (U-HRNet): cloud detection with high-resolution representations for geostationary satellite imagery. *International Journal of Remote Sensing*, 42(9): 3511-3533. <https://doi.org/10.1080/01431161.2021.1875510>
- [20] Zhao, C., Guo, H., Guo, Y., Zhong, S., Luo, H., Fan, J. (2022). TBNet: Two-branch cloud detection network for remote sensing imagery. In 2022 2nd International Conference on Frontiers of Electronics, Information and Computation Technologies (ICFEICT), Wuhan, China, pp. 477-482. <https://doi.org/10.1109/ICFEICT57213.2022.00090>