The International Journal of Engineering and Science (IJES)
|| Volume || 7 || Issue || 10 Ver.IV || Pages || PP 61-67|| 2018 ||
ISSN (e): 2319 – 1813 ISSN (p): 23-19 – 1805

Multi-spectral Cloud Image Cloud Detection Based On Multidimensional Densely Connected Convolutional Neural Network

Bicheng Shi[1],Min Xia[1,2], Hui Cao[1],Wenzhu Shong[1],Jie Wang[1]
[1]*Computer Science, Nanjing University Of Information Science &technology, NanjingJiangsu, Chinese*
[2]*Computer Science, Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing Jiangsu, Chinese*
Corresponding Author: Bicheng Shi

--***ABSTRACT***--

Cloud detection is the prerequisite for multi-spectral satellite cloud image analysis. The traditional cloud detection method cannot well characterize multispectral satellite cloud images, which leads to a low accuracy in cloud detection. Although deep convolutional neural network can extract features effectively, it will have problems such as gradient diffusion low training efficiency,hard optimization, and poor generalization. In order to solve these problems, a multidimensional densely connected convolutional neural network is proposed to realize cloud detection of multi-spectral satellite imagery. Cross layer connection can realize the information flow between all layers in the network, thus reducing the difficulty of convergence caused by the disappearance of the gradient in the training process. The simulation shows that the proposed model can extract the features of cloud images well and improve the accuracy of multi-spectral cloud image detection, it has better generalization performance and higher optimization efficiency.

KEYWORDS;-*Multidimensional densely connected; Convolution neural network; Multi-spectral satellite cloud image; Cloud detection*

Date of Submission: 17-10-2018 Date of acceptance: 03-11-2018

I. INTRODUCTION

With the development of remote sensing satellite technology, remote sensing technology plays a more and more important role in meteorological research and application. Cloud is one of the most important factors in meteorological and climate research. Cloud detection of satellite cloud image has important influence on Surface analysis, weather analysis and environmental assessment, so cloud detection of satellite cloud image becomes more and more important. In the early days, the classification and detection of clouds mainly depended on the selection and judgment by technicians through artificial methods, but with the increasing of remote sensing data, this method became more and more difficult. At present, cloud detection of satellite cloud image has become the core of satellite cloud image application to identify and judge the cloud image effectively and accurately.

Cloud detection standards in cloud images can be based on the criteria of no cloud, partial cloud and complete cloud. Traditional cloud detection methods include threshold method, statistical method, radiation transfer algorithm and comprehensive detection method. The common methods of cloud detection and judgment are statistical method and multi-spectral threshold method[1], clustering method, variational method, wavelet analysis methodand so on. Li Deren et al.[2] put forward a method of using machine learning and multi-feature fusion to realize cloud detection in high-resolution remote sensing images. Luo Peilei et al. extract the features of different depth convolution layers by using the VGG[3], and thenlearn to realize image registration on each convolution layeradaptively.

Traditional cloud detection methods are difficult to extract features and need to extract different features according to different cloud images.There are many limitations in feature extraction and they do not have strong robustness and self-learning ability. In recent years, the deep learning model has achieved good results in various fields. The convolutionneural network can extract the corresponding features according to the target of cloud image detection and classification. The current improved deep neural networks such as VGG ensure the depth of the network by adding a pool layer after several layers rather thanadding a pool layer after each convolution layerimmediately. The experiment proves thatthe accuracy of image classification will be increased when the number of network layers increase, but this growth is very limited, and too many layers will lead to gradient dispersion, so VGG will limit the depth of the network. For the gradient dispersion problem, Kaiming He and others put forward the deep residual network [4]. The network maps the low-level feature map to the upper levelfeature map through "shortcut", and the output of the upper layer becomes the superposition of the bottom mapping and the original output, thus the performance of deeper network is guaranteed by solving the gradient dispersion problem. Compared with the traditional neural network, the above improved methods have great training accuracy and greatgeneralization performance.But the network model is too complex when it train a large number of model parameters. The essential limitation lies in the fact that each convolution layer can only extract features from the previous layer which makes it difficult to make full use of the low-level convolution features and results in the information redundancy of the high-level convolution. In order to solve these problems, multi-dimensional densely connected convolutional neural network is proposed to detect multi-spectral satellite cloud images. Multi-dimensionaldensely connected convolutional neural network can enhance the depth of the network through multi-dimensional dense connection, and effectivelysolve the problemthat the model is too complex and the parameters are too large.The features with complexity in multi-dimension shallow layer are synthetically utilized, and solving the problem that the features of convolution layer can only be transmitted in one dimension. The multi-dimensional transfer of convolution layer features is realized. The experimental results show that the proposed method has less parameter quantity, faster convergence speed and less training time than the previous improved deep convolution neural network model under the premise of ensuring the training accuracy and training effect.

II. MULTIDIMENSIONAL DENSELY CONNECTED CONVOLUTIONAL NEURAL NETWORK

2.1 Structure of densely connected convolution neural network

As shown in figure 1, the details of the multi-dimensional dense connection module. X_1 represents the processed image features. H_l represents composition function. X_l represents the H_l transformed output.

2.1.1 Dense connection mode

The traditional forward connection of convolutional network generally regarded X_l as input and X_{l+1} as output. However, with the deepening of the stacking layer, there will be a more common problem of gradient disappearance.At present, the more common solution to the gradient problem is through the standard initialization and regularization to solve the problem. However, when the network convergence and gradient signal blocking are solved, the model will degenerate. With the increase of the depth of the neural network, the accuracy of the model begins to saturate, and then the saturation will decrease rapidly.Precision degradation shows that the shallow structure can not achieve more layers of deep structure by increasing the number of layers, and the core problem is the added layer is self-mapping and the other layers are copied from the shallow model, so this model does not improve the accuracy. The deep residuals network solves this problem well. The deep residuals network matches the residual map not with a map but by making the additional layer match with a map. The output of the l layer of the deep residual neural network equals the nonlinear transformation of the output of the $l-1$ layer plus the $l-1$ layer, that is:

$$X_l = H_l(X_{l-1}) + X_{l-1}$$

The densely connected convolution neural network [5] is further improved on the basis of the depth residuals network. The output of the l layer is equal to merging the output characteristic map from 1 to $l-1$ layer and then making nonlinear changes，that is:

$$X_l = H_l([x_1, x_2, ..., x_{l-1}])$$

The connection mode of the densely connected convolution neural network enhances the flow of information between the layers of the network, and the dense connection uses a different connection mode. The above formula illustrates this connection more accurately, as can be seen from figure 1. The last layer gets all the previous feature maps, X_1 serve as the initial feature map through the convolution network. Each layer has a nonlinear transformation H_l, and the traditional input of the network H_2 is only X_2, but the input of the dense connection network H_2 has several inputs X_1、 X_2, H_3 has several inputs X_1、 X_2、 X_3, such that each layer is connected to all the other layers, so that each layer has a connection with all the other layers, the l layer and all the previous layers are the inputs of the l layer. Every layer of the network are connected to each other.In this way, the flow of information can be realized better, the training is easier to converge, the training effect can be improved and the training efficiency can be improved.

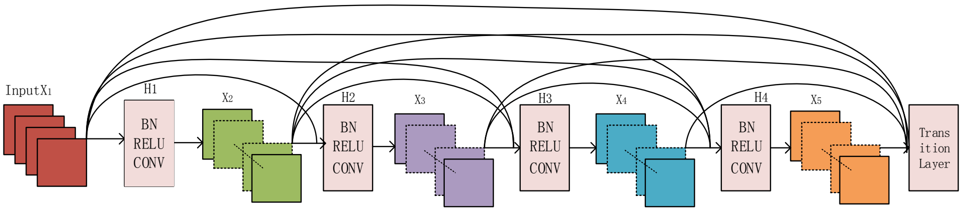


Figure 1 Multi-dimensional Dense Connection Module

2.1.2 Growth rate

After each operation of a composite function, k feature graph is generated, and when the number of layers is l , there will be $l \times k$ feature map.The value of a number of layers represents the amount of information circulating in the network, and the greater of the value k ,the greater of the amount of information circulating in the corresponding network. Correspondingly,the stronger the expression ability of the network, but the value k of the network should not be too large, otherwise, too many parameters of the network will be caused, resulting in too large size and too many parameters of the model.

Table 1 Data of Corresponding Models with Different Growth Rates K

Growth rateK	Modelpara	Training time	Testaccur
K=16	130,253	9670s	88.23%
K=24	255,100	9815s	88.78%
K=32	425,679	9898s	91.36%
K=40	641,750	10741s	91.94%

As shown in Table 1, although the accuracy of the test improves with the increase of the growth rate of k, the model parameters and the training time also increase with the increase of the value of k, in order to quantify the increase of the value of k to limit the value of k, A growth factor is designed:

For every increase in the growth rate k value by 8, the training parameters and training time of the model will increase, and the training accuracy will also increase. The factor means the increase of the accuracy rate after the increase of k value by 8 per time, means the increase of time after the increase of k value by 8 per time, and means the increase of parameters after the increase of k value by 8 per time. In order to achieve the equilibrium effect of each influence factor, the normalization of each factor is done. When k is 24, its growth coefficient is 1.93; when k is 32, its growth coefficient is 15.47; when k is 40, the growth factor is 0.06. The larger of , the better of its performance.when k value is 32, the model can not only achieve better training effect, but also have better network structure.

2.2 Advantages of densely connected convolution neural network

The traditional optimization method adopts the error reverse propagation algorithm. The error signal will only be displayed in the last layer, and then passed from the output layer to the input layer one by one, resulting in the weak supervision of the previous layers, thus making the optimization efficiency lower. The idea of this paper is to add cross-layer connections, and there are connections between any two layers. The l layer contains input from the front $l-1$ layer, turning a one-dimensional linear structure into a network with many cross-layer connections on a plane. When the error signal is calculated forward and back at the last layer, the dense connection can transmit the error signal to any of the previous layers more quickly, and the gradient dissipation will be greatly weakened. The depth of supervision becomes deeper, and the optimization efficiency is improved.

After each layer of the traditional neural network is refined based on the features of the previous layer, the new features are transmitted to the next layer, and the features are abstracted from the lower layer to the high-level feature until the final output. If a certain layer of feature graph is less, the traditional neural network can not optimize the less feature graph.Because the output of any layer of the weighted dense connection network can be used by other layers, the previous features, such as those of the first layer, can be reused at the same time, and then any layer can be reused to the first layer. In this way, only a few features are learned in each layer, which results in a significant reduction in the computational complexity of each layer and a significant reduction in the size of the network parameters.

As shown in figure 2, the network structure of the 1×1 bottleneck layer is used to compress the feature map. Compared with the 3×3 convolution network, the parameters are reduced to $1/9$ of the original, and the parameter scale is greatly reduced.

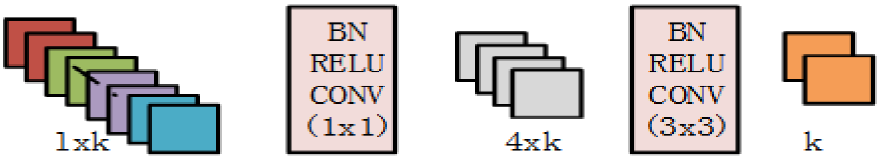


Figure 2 Bottleneck Layer Structure Diagram

As shown in figure 3, if the input has N feature graphs and the output is N feature graphs, then the parameters of this layer of convolution are proportional to the square of N, but for multidimensional weighted densities, its parameters are proportional to $l \times k \times k$, and k is the feature map of each layer, l is the depth of each convolution layer. The biggest difference is that each layer only needs to learn very little features. Generally, k is much smaller than C, so its parameters will be much smaller than the traditional neural network.

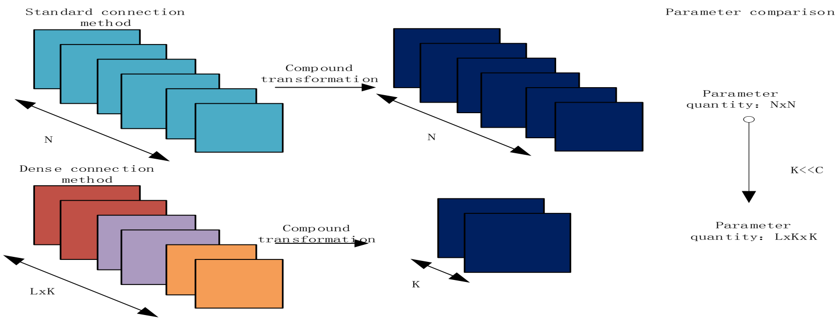


Figure 3 Comparison of Parameters of Different Network Connections

From the point of view of machine learning, there exists a smooth hypothesis.,which is the same as two models when the same training error is achieved, a smoother function of the decision surface is preferred to reduce the occurrence of overfitting, and the generalization performance of the model will be better. Neural network is essentially a composite function, each transformation is equivalent to a composite, multiple recombination will lead to the complexity of the decision function. For traditional neural networks, our classifier function is based on the most complex features, that is, the output based on the last layer. The classifier function is expressed as:

$$Y = w_l H_l(x)$$

Multidimensional densely connected convolution neural network use all the features from simple to complex. They can use not only the most complex features, but also the simple features, so the classifier function obtained by the multi-dimensional densely connected convolution neural network will be the smoothest in the end. The classifier function expression is:

$$Y = w_1 H_1(x) + w_2 H_2(x) + ... + w_l H_l(x)$$

Therefore, the densely connected convolution neural networks also have better performance on small data sets, and have stronger generalization ability than the traditional convolutional neural networks.

2.3 Cloud detection model for multidimensional densely connected convolutional neural networks

The data in this paper come from China Resource Satellite Application Center (http://www.cresda.com/CN), which is one of the three major satellite application centers in China. Cloud image acquisition is mainly from HJ-1A/1B satellite image. For the four visible light channels of HJ-1A/1B, our cloud detection model also increases the dimension of input channel compared with the conventional neural network, and meteorological experts select 9600 thick cloud samples, 9600 cloudless samples and 9600 thin cloud samples from the original cloud map. 28,800 samples are taken as data sets, 24, 000 samples are taken as training sets, and the remaining samples of each class are taken as testing sets. The pixel size of each sample is 28×28.

In this paper, there are 4 dense connection modules in the 121 layer cloud detection model, and the size of the feature diagrams in each dense connection module remains uniform, and only 32 feature diagrams are added to each layer of each dense connection module. The first dense module has 6 layers, the second module has 12 layers, the third module has 24 layers, and the fourth module has 16 layers. Finally, each feature graph is compressed by convolution and pool operation. All the feature graphs form one-dimensional data, connect with three outputs, and finally output the probability of each category.

III. RESULT VIEW

3.1 Cloud Detection Effect Graphs

Figure 4 shows the effect of cloud detection under different methods.Fig. 4 (a) is the original image of satellite cloud map, Fig. 4 (b) is the effect graph of BP-Net, Fig. 4 (c) is the effect graph of CNN, Fig. 4 (d) is the effect graph of VGG, Fig. 4 (e)is the effect graph of Resnet, Fig. 4 (f) is the effect graph of M-Densenet.

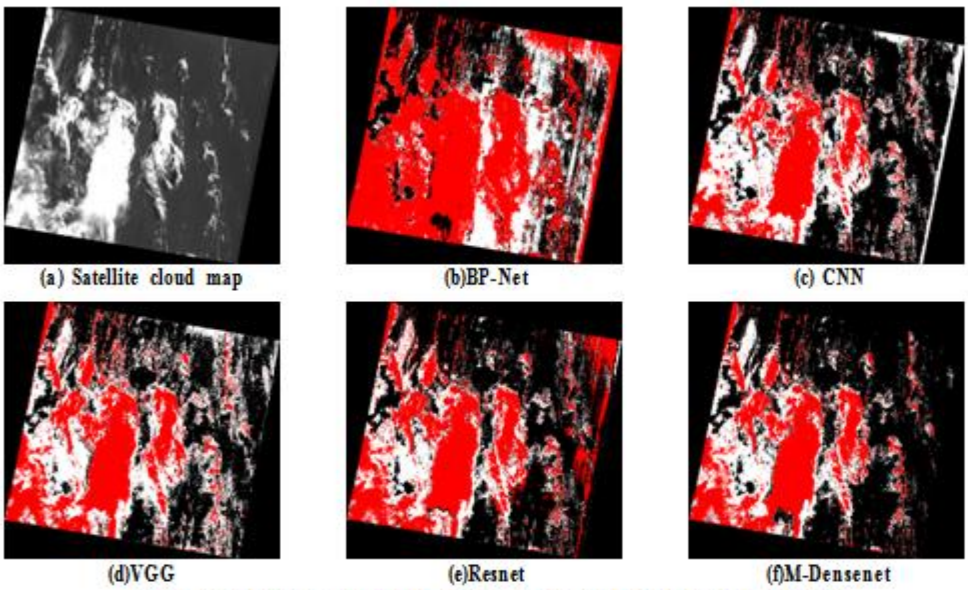


Figure 4 Comparison of Cloud Detection Effect Graphs by Different Methods

The thick cloud area is represented by red, the thin cloud area by white, and the cloudless region by black. By analyzing the effect of each detection in the graph, we can get the following conclusion: figure 4, the red part of (b) is too much, The black part is too few, BP-Net is a kind of multi-layer feedforward neural network trained according to the error back propagation algorithm [6]. Its biggest characteristic is that it has very strong nonlinear mapping ability and flexible network structure, but the convergence speed is slow. And the generalization ability is poor, the thick cloud in the map exists more serious overdetection. Figure 4: there is a white edge overdetection in (c). Convolution neural network is a kind of convolutional neural network. The feedforward neural network used in large scale image processing, because of the existence of local receptive field [7], parameter sharing , pool and so on, greatly reduces the parameters of the model, but in the depth of the model, With the increase of the depth of the model, the gradient will disappear or the gradient will explode [8]. The generalization ability of the convolution neural network in this paper still has some problems, and the edge of the graph will be misdetected as the thin cloud. Figure 4. In (d), there are also thin clouds in the right half of the picture. The improvement of VGG convolution neural network is characterized by deepening the number of layers of the network and reducing the size of the filter. The performance of the algorithm is optimized, but it inevitably results in too much parameter. Figure 4: there is overdetection of thick cloud in the edge of (e). The improvement of Resnet convolution neural network can further increase the depth of the network and greatly improve the training accuracy by introducing a "shortcut connection"[9]that can skip one or more layers. However, there are still some problems such as low training efficiency and weak generalization ability. Figure 4 (f) meets our requirements fairly well and optimizes the efficiency through thetransfer of low, medium, and high multilayer features. And generalization ability has been greatly improved,model distinguish thick cloud and thin cloud better, andhas achieved better resultsin the distinction between thin cloud and no cloud .

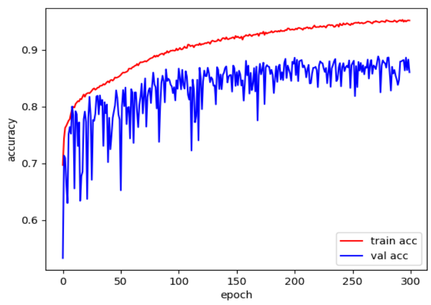
In order to make a more effective quantitative analysis of the results, this paper uses BP-Net, CNN, VGG, ResNet, M-Densenet to compare the results, as shown in Table 2:

Table 2 average accuracy of satellite cloud images

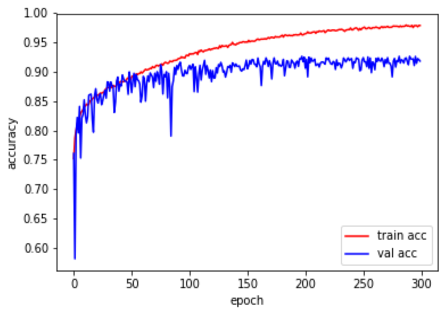
detection Algorithm	accuracy rate
BP	72.46%
LeNet	82.45%
VGG	85.34%
Resnet	86.21%
M-Densenet	91.36%

The data in Table 2 are based on the experimental results of 200 satellite original cloud images from different regions. Because the traditional deep learning model such as back propagation neural network (BP), convolutional neural network (LeNet) is too shallow to have a good feature expression ability so it does not have a strong generalization ability. Secondly, VGG convolution neural network (VGG), ResNet convolutional neural network (ResNet) increases the depth of convolution layer on the basis of traditional depth learning model and improves the performance of the model. However, with the increase of the number of layers, there will be gradient problems of varying degrees, and the lack of effective use of low-level features can only continuously extract high-level features from low-level features. It is difficult to obtain better results in the deeper network model because of the low utilization rate of low-level features. Due to the deep depth of the multi-dimensional weighted dense connection network and the connection between each layer and all the previous layers, the low, middle and high level features are utilized efficiently, which makes the convolutional neural network achieve better classification effect in the deeper case. And the generalization ability is also improved.

3.2 Analysis of simulation result



(a) Training and Validation accuracy of Resnet



(b) Training and Validation accuracy of M-Densenet

The trend of training accuracy and validation accuracy on remote sensing cloud image dataset is shown in Fig. 7.It can be seen from the graph that the convergence speed of themulti-dimensional densely connected convolutional network is obviously higher than that of the residual neural network. In terms of accuracy,The training accuracy of the multi-dimensional densely connected convolutional network is close to that of multi-dimensional densely connectedconvolution neural network, while the accuracy rate of depth residual convolution neural network is close to that of iterative step. The training accuracy of multi-dimensional weighted dense convolutional neural network is higher than that of multi-dimensional weighted dense convolution neural network. It is obviously superior to the deep residual convolution neural network. The effect of verifying accuracy is also clear. Show advantage. The results show that the training effect of convolution neural network based on multidimensional dense connection is better than that of deep residual convolution neural network.

The hardware configuration of the experimental environment is Intel Core i7-6700k eight-core processor, GeForceGTX980 4G independent graphics card, 16G memory, 1T hard disk, and the software simulation framework is Keras framework. In order to better verify the superiority of this model, the parameters and training time of the improved depth model are calculated, as shown in Table 3:

Table 3 comparison of improved depth model parameters

detection model	model parameter	training time
VGG-19	51,516,547	10023.85s
ResNet	58,457,988	17726.79s
Densenet	425,679	9898s

Combined with Table 3, we can find that the improved algorithm improves the complexity of the algorithm, both in space and in time. The spatial complexity of the convolution neural network is mainly determined by the size of the convolution kernel K, the number of channels C and the depth of the network D.

$$Space \sim O\left(\sum_{l=1}^{D} K_l^2 \bullet C_{l-1} \bullet C_l\right)$$

The improved convolution neural network algorithm in this paper is similar to other improved convolution neural networks, although the network depth D is close to the size of the convolution kernel K. However, the proposed algorithm only adds a small number of channels C_l to each layer, which greatly reduces the parameters of the model compared with other convolution neural networks. The time complexity of the same convolution neural network is mainly determined by the size of the output feature graph of each convolution kernel M, the size of the convolution kernel K, the number of channels, and the depth of the network D.

$$Time \sim O\left(\sum_{l=1}^{D} M_l^2 K_l^2 \bullet C_{l-1} \bullet C_l\right)$$

Although the improved convolution neural network is improved in depth and generalization ability, it can not solve the problem that every layer of convolutional neural network training requires a large number of channels, resulting in an increase in time complexity. The improved method in this paper is far less than other improved deep learning models on the premise of ensuring accuracy and training effect. The computational resources are greatly saved, and the speed is improved compared with other models, which solves the problem of too many parameters of previously improved convolution neural networks.

IV. CONCLUSION

In this paper, the multi-dimensional densely connected convolutional networkcombine with cloud image detection. By adding dimension, it can train the multi-spectral cloud image data set. In this paper, the multi-dimensional densely connected convolutional network, with its unique weighted connection and cross-layer dense connection, enables the gradient signal to be effectively transmitted between layers. It solves the problem that the traditional neural network is difficult to train due to the disappearance of the gradient and the explosion of the gradient with the increase of the number of layers, improves the depth of the convolutional neural networkeffectively, and improves the accuracy and generalization performance greatly. And solved the previous changes. The features of progressive convolution neural networks can only be transmitted in one dimension, which realizes the multidimensional transmission of features of low, middle and high levels. The multi-dimensional weighted densely connected convolutional network has good performance in both the overall effect map and the accuracy trend graph. However, there are some improvements in the multi-dimensional weighted densely connected convolutional network in this paper. The model will be partially overfitted, so there will be partial misinformation in the cloud map. Therefore, the next step is to further optimize the network, solve the problem of partial overfitting, and further improve the generalization ability.

REFERENCE

[1]. Paola J D, Schowengerdt R A.A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery[J].International Journal of Remote Sensing, 1993, 16(16):3033-3058

[2]. Bai T, Li D, Sun K, et al. Cloud Detection for High-Resolution Satellite Imagry Using Machine Learning and Multi-Feature Fusion[J]. Remote Sensing, 2016, 8(9):715.

[3]. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.

[4]. He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. 2015:770-778.

[5]. Huang G, Liu Z, Maaten L V D, et al.Densely Connected Convolutional Networks[C].IEEE Conference on Computer Vision and Pattern Recognition, 2017, :2261-2269

[6]. Santos R B, Rupp M, Bonzi S J, et al.Comparison Between Multilayer Feedforward Neural Networks and a Radial Basis Function Network to Detect and Locate Leaks in Pipelines Transporting Gas[C]// International Conference on Chemical and Process Engineering. 2013:1375-1380.

[7]. Huang G B, Bai Z, Chi M V.Local Receptive Fields Based Extreme Learning Machine[J].IEEE Computational Intelligence Magazine, 2015, 10(2):18-29.

[8]. Sudicky E A, Cherry J A, Frind E O. Migration of contaminants in groundwater at a landfill: A case study : 4. A natural-gradient dispersion test[J]. Journal of Hydrology, 1983, 63(1–2):81-108.

[9]. Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning[J]. 2016.

Bicheng Shi "Multi-spectral Cloud Image Cloud Detection Based On Multidimensional Densely Connected Convolutional Neural Network "The International Journal of Engineering and Science (IJES),), 7.10 (2018): 61-67