Research on Cloud Recognition Technology Based on Transfer Learning

Chunyao Fang1,2, Kebin Jia*,1,2, Pengyu Liu1,2, Liang Zhang1,2
1 Faculty of Information Technology, Beijing University of Technology, Beijing, 100124 China
2 Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing University of Technology, Beijing, 100124 China

*E-mail: kebinj@bjut.edu.cn

Abstract—The cloud is an important part of the earth's thermodynamic balance and water and air cycle. At present, abundant achievements have been made in the research of satellite cloud image, while the recognition of ground-based cloud image has always been a difficulty in the field of pattern recognition, and the achievements are relatively limited. In this paper, based on the ground-based cloud map data set provided by standard weather stations, after data enhancement, 5 network models were trained by means of fine-tuning network parameters and freezing weights of different network layers, and 5 network migration configurations were used on the enhanced data set. Experimental results show that the fine-tuned Densenet network model is more suitable for this project, and the recognition accuracy can reach 96.55%.

I. INTRODUCTION

Clouds are an important part of the earth's thermal balance and water and air cycle. Their morphology reflects the stability of atmospheric movement and weather characteristics, and is one of the main characteristics that predict future weather changes. Clouds can be divided into ten genera according to their appearance characteristics, structural characteristics and cloud bottom height, etc. Accurately identifying each cloud shape can help us effectively improve the accuracy of weather forecasting, the effectiveness of climate model prediction and the understanding of global climate change, which is of great significance [1]. At present most of the stations still rely on cloud meteorological observer artificial visual recognition, but artificial recognition are susceptible to the spotter mood experience, observing the influence of subjective factors, such as in addition, because of the cloud of rigid structure, unpredictable and artificial observation workload is huge, its shape is difficult to guarantee the continuous observation for a long time, restricted the cloud recognition accuracy [2]. In recent years, with the rapid development of digital image processing and pattern recognition technology, automatic cloud recognition technology based on cloud image has become a hotspot in meteorology [3-5].

II. RELATED WORK

Automatic cloud identification technology can be divided into two aspects: satellite cloud image and ground-based cloud image [6]. Compared with satellite cloud images, ground-based cloud images have higher spatial resolution and richer local cloud information [7]. Therefore, automatic cloud identification based on ground-based cloud images is attracting more and more attention. Scholars at domestic and abroad have carried out relevant researches and achieved certain results. Han wenyu et al. used the sparsity of cloud image gray scale to identify 5 types of cloud images, with an average accuracy of 82.8%. However, the reconstruction performance of this algorithm was poor, and only the typical cloud images with high discrimination could be recognized, and the similar cloud images such as stratified clouds and cumulonimbus clouds could not be distinguished [8]. Li chenxi et al. conducted scaling analysis on the gray scale data of cloud images by extended self-similarity (ESS) model. 150 cloud images of 5 categories were identified by using scaling features of different cloud systems, with an accuracy rate of nearly 90%. However, due to the limited number of test samples, the robustness of the algorithm could not be verified [9]. Zhang chi et al. proposed a cloud image recognition method based on visible light and infrared image information fusion, and identified 5 cloud images, with an average accuracy of 82%. Because this algorithm requires comprehensive consideration of the whole sky cloud image and infrared cloud image, the recognition conditions are harsh and the applicability is not high [10]. Since the traditional image recognition algorithm cannot adapt to the changeable characteristics of cloud, and the recognition types are few and the accuracy is low, scholars turn the center of gravity of cloud recognition to the deep learning method. Zhao lianglei et al. used convolutional neural network for cloud recognition, proving that convolutional neural network has higher detection accuracy than traditional classification methods [11]. Zhong Zhang et al. proposed the transmission depth local binary mode (TDLBP) and weighted metric learning (WML), combined with convolutional neural network, to identify cloud images collected from different perspectives with an accuracy rate of nearly 80% [12]. Jinglin Zhang, etc. This paper proposes a new CloudNet convolution neural network model, the identification of 11 class cloud, with 88% accuracy, but the experiment data set of data are characterized by obvious distinguish degrees cloud image, so
the identification accuracy is higher, in addition, due to the small data set, contains only 2543 piece of cloud, there may be a fitting accuracy caused by artificially high [13]. With the rapid development of artificial intelligence, all kinds of image recognition technologies have become mature [14]. However, there are not many studies on cloud image recognition based on ground-based cloud image. Clouds are divided into 10 genera. However, due to the varied shapes and high complexity of clouds, the existing researches can only ensure the identification of several typical genera. The recognition accuracy is not high and needs to be further improved [15,16].

In conclusion, the research and design of a precise and automatic identification technology for a variety of clouds is of great significance to improve the accuracy of weather forecast, flight support, analysis of regional climate characteristics, and has great application value. Considering the convolution the outstanding performance of the neural network in the field of image recognition, this paper provided by the standard meteorological stations in the same point of view the collected data to enhance the resolution of the same class 11 cloud, augmented dataset, and then based on the parameters of the migration method of study, the five classical convolution neural network model parameters fine-tuning and different weights of the network layer is frozen, contrast test results, select the optimal network, has realized the accurate identification of 11 class cloud, solves the existing research results, select the optimal network, has realized the accurate different weights of the network layer is frozen, contrast test convolution neural network model parameters fine-tuning and view the collected data to enhance the resolution of the same point of view the collected data to enhance the resolution of the same cloud group, middle cloud group and high cloud group [17].

III. THE OVERALL STRUCTURE OF CLOUD RECOGNITION ALGORITHM

Considering the non-rigid feature of cloud, the recognition algorithm based on convolutional neural network has higher applicability and recognition accuracy than traditional image recognition algorithm. Therefore, this paper studies a cloud recognition technology based on transfer learning method, and adjusts several existing convolutional neural networks to get the most suitable cloud recognition network. The overall structure of the algorithm is shown in Fig. 1.

Firstly, the original data set provided by standard meteorological stations is denoised. Considering the sensitivity of convolutional neural network recognition algorithm to data volume, this paper performs data enhancement operation on the original data set and expands the cloud image data to five times of the original data. Secondly, in network training, the transfer learning method based on model is adopted to replace classifiers and freeze weights of network layer for the classical convolutional neural network respectively. The optimal model is selected by comparing the accuracy of cross-validation in the training process. Finally, the cloud image to be identified is input into the pre-training network model to obtain the classification results.

A. Classification And Pretreatment Of Ground-based Cloud Maps

Cloud shape refers to the physical characteristics of clouds. According to the shape, composition and cause of formation of clouds, clouds are divided into Cumulus, Cumulonimbus, Stratocumulus, Stratus, Nimbostratus, Altostratus, Altocumulus, Cirrus, Cirrostratus and Cirrocumulus. According to the height of the ten genera, they can be divided into three groups: low cloud group, middle cloud group and high cloud group [17].

The classification and characteristics of cloud genera are shown in Table I.

B. Pretreatment

(1) Data source and Processing.

Due to the shortcomings of the existing ground-based cloud image recognition research, such as small open source data set, lack of cloud genus and different picture quality, this paper uses the standard cloud map provided by Huayunshengda (Beijing) meteorological technology limited liability company. This database is composed of 11 cloud maps (10 clouds and no clouds), with a total data amount of 9517 ground-based cloud maps with a resolution of 1358×1358. Considering that the data set provided by Huayun is small, and the data among various categories are extremely unbalanced, it is easy to overfit. In this paper, horizontal inversion, vertical inversion, fixed Angle rotation, translation and random gaussian noise processing are adopted to conduct non-directional data enhancement for cloud images of different categories. The data volume of the enhanced data set is nearly 5 times that of the original data set. The composition of original data set and enhanced data set is shown in Table II.

(2) Data set comparison.

As can be seen from Table III and Fig. 2, the existing open source cloud map data sets are few, with low resolution and different cloud image quality. Although CCSN data set contains all cloud genera, due to its data collection from different angles, the images contain many interferences (trees, houses, etc.), resulting in different picture quality. HUST and SWIMCAT data set, however, contain incomplete cloud
genus categories and are not classified according to the standard definition of cloud genus, with low reference value. The cloud pictures adopted in this paper is provided by professional weather stations, and the pictures are collected from the same Angle. The cloud genus is divided according to industry norms, and the quality is uniform. Compared with other existing data sets, it has a higher reference value.

IV. THE CONVOLUTIONAL NEURAL NETWORK AND TRANSFER LEARNING

A. The Convolutional Neural Network

Convolutional neural network (CNN) is a deep feedforward neural network based on convolution operation. Convolution

<table>
<thead>
<tr>
<th>Cloud class</th>
<th>Scientific name</th>
<th>Abbreviation</th>
<th>Cloud genus</th>
<th>Cloud characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low cloud group</td>
<td>Cumulus</td>
<td>Cu</td>
<td>Having an upward projection of a circular arch; clouds similar in size to fists; margins clear.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cumulonimbus</td>
<td>Cb</td>
<td>Clouds are thick and broccoli-like; edges are blurred.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stratocumulus</td>
<td>Sc</td>
<td>Clouds are generally fist-sized and loosely distributed, clustered, traveling and wavy, often grey or gray-white.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stratus</td>
<td>St</td>
<td>Clouds lay evenly; cover a large area, almost all over the sky; mostly grey.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nimbostratus</td>
<td>Ns</td>
<td>Clouds are low and amorphous; often covered with the sky and completely obscured the sun and moon; clouds are fluffy and dark grey.</td>
<td></td>
</tr>
<tr>
<td>Middle cloud group</td>
<td>Altosstratus</td>
<td>As</td>
<td>Clouds are thicker and covered with the sky; the sun passes through almost no halo; clouds often have striped structures and are grayish-white or grayish-blue.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Altocumulus</td>
<td>Ac</td>
<td>Clouds are dark grey; the outline of the Sun-Moon is not clear; clouds are oval, tile-shaped, fish scales or water wavy distribution.</td>
<td></td>
</tr>
<tr>
<td>High cloud group</td>
<td>Cirrus</td>
<td>Ci</td>
<td>Thin and transparent; white and shiny; the clouds are filamentous and horsetail-like.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cirrostratus</td>
<td>Cs</td>
<td>The bottom of the cloud has a filamentous structure; the cloud body is thin enough to pass through the sun and the moon; and there is a distinct halo under the sun's illumination.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cirrocumulus</td>
<td>Cc</td>
<td>Clouds are very small, white and shiny; they are thin white scales; they are often arranged in rows and in groups.</td>
<td></td>
</tr>
</tbody>
</table>

TABLE II

<table>
<thead>
<tr>
<th>Cloud type</th>
<th>Ae</th>
<th>As</th>
<th>Cc</th>
<th>Ci</th>
<th>Cs</th>
<th>Cu</th>
<th>Sc</th>
<th>Cb</th>
<th>Ns</th>
<th>St</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data volume</td>
<td>376</td>
<td>1292</td>
<td>217</td>
<td>1025</td>
<td>1319</td>
<td>775</td>
<td>116</td>
<td>103</td>
<td>132</td>
<td>3804</td>
</tr>
<tr>
<td>Enhanced data volume</td>
<td>5114</td>
<td>5047</td>
<td>5290</td>
<td>5607</td>
<td>5292</td>
<td>5526</td>
<td>5232</td>
<td>574</td>
<td>735</td>
<td>791</td>
</tr>
</tbody>
</table>

Fig.2 (a) Cloud samples from CCSN, (b) cloud samples from HUST, (c) cloud samples from SWIMCAT and (d) the data set used in this paper.
is a special linear operation. Convolution network refers to a neural network that uses convolution operation to replace the general matrix multiplication operation at least one layer in the network. CNN has two major features: first, the network structure contains at least one convolutional layer to extract features; second, its convolutional layer works through weight sharing, which reduces the complexity of the network.

B. Transfer Learning

Training a network model for a certain task requires a large amount of data and computing power, and building a large and accurate data set requires a lot of time, manpower and material resources, which is difficult to achieve in a short time. How to make use of the existing mature network structure to carry out targeted retraining for specific tasks has become a very meaningful problem, therefore, transfer learning has attracted the attention of many researchers.

The core of transfer learning is to find the similarity between existing knowledge and new knowledge, so that existing knowledge can be used to assist the rapid learning of new knowledge. According to different learning styles, transfer learning can be divided into sample-based transfer, feature-based transfer, model-based transfer and relation-based transfer. In convolutional neural network, training a brand-new network model requires millions of parameters, which is time-consuming and difficult. Compared with the traditional network training method, the advantage of migration learning is that it does not need a large number of data samples. By fine-tuning the network model of similar tasks, model reuse can be achieved without retraining, thus greatly reducing the operation time.

For the task of cloud recognition, due to the high similarity of some clouds, the effect of the methods based on samples, features and relationships is not obvious, so this paper chooses the transfer learning method based on model. The key steps can be divided into three steps: model selection, model training and detection accuracy. These three steps are described in detail below:

1) Model selection

In this paper, transfer learning and training samples are carried out based on 5 networks with different structures. The networks adopted are as follows:

1) VGG16.

The VGG16 network consists of 16 layers requiring training parameters (excluding the pooling layer), namely 13 convolution layers and 3 full connection layers. The outstanding features of this network structure are as follows: the convolution layer adopts the same convolution core parameters, the size of convolution core is 3x3; the pooling layer adopts the same pooling core parameters, the size is 2x2; the model is composed of several convolution layers and stacks of pooling layers, so it is easy to form a deeper network structure.

2) VGG19.

The VGG19 network includes 19 layers requiring training parameters, namely 16 convolution layers and 3 full connection layers. In conclusion, the advantages of VGG network model can be summarized as: small filters, sinking networks.

3) Inception-V3.

The Inception-V3 network is characterized by splitting two large two dimensional convolution into two smaller one dimensional convolution. On the one hand, overfitting is avoided by fewer parameters; on the other hand, a layer of nonlinear extended model expression capability is added. In addition, the results of asymmetric convolution structure decomposition (such as 3x3 decomposition into 1x3 and 3x1) can handle more abundant spatial features and increase feature diversity compared with the results of symmetric decomposition into several identical small convolution kernel.

4) ResNet152.

The main idea of ResNet Network is to add a direct connection channel, namely Highway Network. Highway Network allows to retain a certain proportion of the output of the previous Network layer, and allows the original input information to be directly transmitted to the later layer, as shown in Fig. 3. Its network layer number is 152, won the first prize in ILSVRC2015 competition, the error rate in top-5 is 3.57%, and the number of participants is lower than VGG network, the effect is outstanding. The ResNet residual learning module is shown in Fig. 3.

5) DenseNet201.

The basic idea of DenseNet is the same as ResNet, except that it establishes a dense connection between all the previous layers and the later layers, that is, each layer accepts all the previous layers as its additional input. It is different from ResNet in that it uses its features re-use by connecting them to the channel. Therefore, it has better performance than ResNet in the case of fewer parameters and computational costs, and solves the problem of gradient disappearance caused by deepening network layers.

![Fig. 3 ResNet residuals learning module.](image)

(2) Model training

1) The experimental setup.

This experiment was carried out on Ubuntu18.04 operating system, using PyTorch framework and python3.6.6 programming language. The cross entropy function was selected as the objective optimization function in the training process. In this paper, fine-tuning training was carried out for each network model after the classifier was replaced in the experiment, and different neural network layers were thawed out. The best training results of each model during fine-tuning were recorded for data analysis and comparison.
2) The experimental means.

In this paper, classifier replacement and model network layer weight freezing are adopted for experiments. Among them, the classifier replacement form is to conduct experiments on the original data set and the enhanced data set respectively, and conduct the training results comparison of the two data sets. The parametric freezing method only applies to the enhanced data sets, freezes the weights of different networks to different degrees, and compares the training effects.

a. Classifier replacement.

Considering the network to identify the target of the original model and to identify the 11 class cloud image, therefore, when using the migration study in this paper, first download has completed training of network weights, each layer of network training, the last a fully connected network layer replacement for 11 kinds of cloud classifier to identify, the weight of this operation is equivalent to the source model as a preliminary training parameter, to the entire network for training.

When training parameters are set, the size of each batch of data is set as 32, the learning rate is 0.001, and the number of iterations is set as 100 by default. However, early stop mechanism is added in the training process, and the training will be stopped when the verification loss is not reduced for 5 consecutive iterations. The training effect of classifier replacement method on two data sets is shown in Table IV. When freezing the weight of network layer, all the parameters of all layers except the last layer are frozen (that is, the setting parameters cannot be updated), and then they are thawed forward layer by layer. Five classical network models were thawed at different levels, and the best accuracy rate of thawed forward layer by layer. Five classical network models, and each layer was carried out by using five network models, and each

b. Model network layer weight freezing

When freezing the weight of network layer, all the parameters of all layers except the last layer are frozen (that is, the setting parameters cannot be updated), and then they are thawed forward layer by layer. Five classical network models were thawed at different levels, and the best accuracy rate of each training result was recorded. The recognition accuracy rate of training set and test set was used as evaluation indexes to evaluate the effects of the five models. The training effect is shown in Table V.

(3) Detection accuracy

Accuracy is a commonly used image classification and information retrieval of objective evaluation criteria, defined as (1):

\[
\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

that which TP (True Positive) is model classification is correct sample quantity, TN (TrueNegative) is model classification number of Negative samples correctly, FN (False Negative) for the model classification error is sample quantity, FP (False Positive) for the model classification error of Negative samples. In the training process, the training samples will be tested once for each iteration, and the next iteration will be decided according to the trend of accuracy. Finally, the new model is used to detect accuracy in the verification center.

V. EXPERIMENTAL RESULTS AND ANALYSIS

A. Classifier Replacement.

<table>
<thead>
<tr>
<th>The network model</th>
<th>Original data set</th>
<th>Enhanced data set</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG16</td>
<td>86.52%</td>
<td>86.3%</td>
</tr>
<tr>
<td>VGG19</td>
<td>84.33%</td>
<td>84.7%</td>
</tr>
<tr>
<td>Inception-V3</td>
<td>84.0%</td>
<td>87.2%</td>
</tr>
<tr>
<td>ResNet152</td>
<td>88.21%</td>
<td>90.6%</td>
</tr>
<tr>
<td>DenseNet201</td>
<td>89.1%</td>
<td>92.8%</td>
</tr>
</tbody>
</table>

Considering the high sensitivity of convolutional neural network to data volume, this paper adopts classifier replacement method for experiments based on data amplified and equalized data set and original data set respectively. As can be seen from the experimental results, compared with the recognition effect of the original data set, the enhanced data set has higher recognition accuracy under different networks, which proves the necessity of this paper to build a data balanced and uniform cloud map data set with high data volume, laying a foundation for subsequent experiments.

B. Model Network Layer Weight Freezing

<table>
<thead>
<tr>
<th>The network model</th>
<th>Thawed layer</th>
<th>Testing set accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG16</td>
<td>86.52%</td>
<td></td>
</tr>
<tr>
<td>VGG19</td>
<td>86.33%</td>
<td></td>
</tr>
<tr>
<td>Inception-V3</td>
<td>84.0%</td>
<td></td>
</tr>
<tr>
<td>ResNet152</td>
<td>96.15%</td>
<td></td>
</tr>
<tr>
<td>DenseNet201</td>
<td>96.55%</td>
<td></td>
</tr>
</tbody>
</table>

In this paper, the weight freezing experiment of network layer was carried out by using five network models, and each layer was thawed forward to five layers. The accuracy of each test set was recorded, and the result of thawing layer with the highest accuracy was selected as the reference of network model selection. As can be seen from the experimental results in Table V, when DenseNet201 thaws two layers forward, it has higher recognition accuracy than the optimal identification results of other networks under different thawing layers, which proves that the method and network selected in this paper have higher advantages than existing cloud recognition algorithms.

<table>
<thead>
<tr>
<th>The experimental method</th>
<th>Number of recognized species</th>
<th>Data volume</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud recognition algorithm based on ESS model [9]</td>
<td>5</td>
<td>150</td>
<td>90%</td>
</tr>
<tr>
<td>An algorithm of ground cloud image recognition based on selective ensemble neural network based on K mean value [18]</td>
<td>4</td>
<td>800</td>
<td>94%</td>
</tr>
<tr>
<td>Cloud recognition algorithm based on infrared information fusion [10]</td>
<td>5</td>
<td>552</td>
<td>82%</td>
</tr>
<tr>
<td>CloudNet [13]</td>
<td>11</td>
<td>2543</td>
<td>88%</td>
</tr>
<tr>
<td>TDLBP+WML [12]</td>
<td>7</td>
<td>3533</td>
<td>78.52%</td>
</tr>
<tr>
<td>The algorithm in this paper (DenseNet unfreezing 2 layers)</td>
<td>11</td>
<td>44968</td>
<td>96.55%</td>
</tr>
</tbody>
</table>
The optimal algorithm in this paper (DenseNet model layer weight freezing method) is compared with recent research results, as shown in Table VI. It can be seen that the algorithm in this paper is not only of low complexity, but also significantly superior to other algorithms in recognition category and accuracy. In addition, due to the sufficient amount and uniform quality of data used in this document, the incidence of overfitting during model training can be effectively reduced, and the robustness and applicability of the algorithm can be guaranteed.

VI. CONCLUSION

In the experimental part, two methods, classifier replacement and model parameter fine-tuning, are used to carry out different experiments. The classifier replacement method only carries out experiments on the original data set and the enhanced data set. The experimental results verify the sensitivity of convolutional neural network to data volume. Using the amplified data set for network training can achieve significantly better results, proving the necessity of the amplified data set. The parameters of 5 classical network models were fine-tuned, and the optimal results were selected for comparison. The experimental results highlighted the significant advantages of DenseNet201 in cloud recognition direction. When thawing 2 layers, the recognition effect reached the best, and the accuracy of test set was significantly better than that of other networks. Comparing this algorithm with other algorithms in recent years, it can be seen that this algorithm has significant advantages in algorithm complexity, recognition type and accuracy. For the later work, considering the disadvantage that the training time of network model increases explosively with the increase of thawing layers, it is proposed to build a network structure specially used for cloud identification by referring to the characteristics of residual module of DenseNet201. The ab initio training network can ensure the accuracy of cloud identification under the premise of effectively reducing the training time. In addition, due to the small amount of data of cumulonimbus cloud, rain layer cloud and layer cloud, it is still difficult to maintain balance with other cloud map data even after data enhancement. We plan to use network crawler technology to obtain more data in the later work, and conduct the next experiment under the guarantee of data balance.

ACKNOWLEDGMENT

This paper is supported by the following funds: Developments of major scientific instruments and equipment (2018YFF01010100), Nation natural science foundation of China(61672064), Beijing natural science foundation project(4172001) and Advanced information network Beijing laboratory(PXM2019_014204_500029).

REFERENCES


