



## Article

# A Cloud Classification Method Based on a Convolutional Neural Network for FY-4A Satellites

Yuhang Jiang <sup>1</sup>, Wei Cheng <sup>2,\*</sup>, Feng Gao <sup>1,3</sup> , Shaoqing Zhang <sup>1,4,5,6,7</sup>, Shudong Wang <sup>8</sup>, Chang Liu <sup>1,9</sup> and Juanjuan Liu <sup>10</sup> 

<sup>1</sup> College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China; jiangyh97@hrbeu.edu.cn (Y.J.); gaofeng19@hrbeu.edu.cn (F.G.); szhang@ouc.edu.cn (S.Z.); liuchang407@hrbeu.edu.cn (C.L.)

<sup>2</sup> Beijing Institute of Applied Meteorology, Beijing 100029, China

<sup>3</sup> Qingdao Innovation and Development Center, Harbin Engineering University, Qingdao 266400, China

<sup>4</sup> The College of Ocean and Atmosphere, Ocean University of China, Qingdao 266100, China

<sup>5</sup> Key Laboratory of Physical Oceanography, MOE, Institute for Advanced Ocean Study, Frontiers Science Center for Deep Ocean Multispheres and Earth System (DOMES), Ocean University of China, Qingdao 266100, China

<sup>6</sup> Ocean Dynamics and Climate Function Lab/Pilot National Laboratory for Marine Science and Technology (QNL), Qingdao 266237, China

<sup>7</sup> International Laboratory for High-Resolution Earth System Prediction (IHESP), Qingdao 266000, China

<sup>8</sup> Public Meteorological Service Center, China Meteorological Administration, Beijing 100081, China; wangsd@cma.gov.cn

<sup>9</sup> Qingdao Hatran Ocean Intelligence Technology Co., Ltd., Qingdao 266400, China

<sup>10</sup> State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China; ljxgg@mail.iap.ac.cn

\* Correspondence: chengw@mail.iap.ac.cn



**Citation:** Jiang, Y.; Cheng, W.; Gao, F.; Zhang, S.; Wang, S.; Liu, C.; Liu, J. A Cloud Classification Method Based on a Convolutional Neural Network for FY-4A Satellites. *Remote Sens.* **2022**, *14*, 2314. <https://doi.org/10.3390/rs14102314>

Academic Editor: Vijay Natraj

Received: 21 March 2022

Accepted: 9 May 2022

Published: 11 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** The study of cloud types is critical for understanding atmospheric motions and climate predictions; for example, accurately classified cloud products help improve meteorological predicting accuracies. However, the current satellite cloud classification methods generally analyze the threshold change in a single pixel and do not consider the relationship between the surrounding pixels. The classification development relies heavily on human recourses and does not fully utilize the data-driven advantages of computer models. Here, a new intelligent cloud classification method based on the U-Net network (CLP-CNN) is developed to obtain more accurate, higher frequency, and larger coverage cloud classification products. The experimental results show that the CLP-CNN network can complete a cloud classification task of  $800 \times 800$  pixels in 0.9 s. The classification area covers most of China, and the classification task only needs to use the original L1-level data, which can meet the requirements of a real-time operation. With the Himawari-8 CLTYPE product and the CloudSat 2B-CLDCLASS product as the test comparison target, the CLP-CNN network results match the Himawari-8 product highly, by 84.4%. The probability of detection (POD) is greater than 0.83 for clear skies, deep-convection, and Cirrus–Stratus type clouds. The probability of detection (POD) and accuracy are improved compared with other deep learning methods.

**Keywords:** cloud classification; FY-4A; AGRI; deep learning; geosynchronous satellites

## 1. Introduction

Clouds have a very important role in weather systems. Approximately 70% of the global cloud volume can significantly affect the hydrological cycle and radiation budget of the global atmosphere [1,2]. Different cloud types have different radiative effects on the surface–atmosphere system. Accurate and automatic cloud detection and classification can have a huge impact on many climatic, hydrological, and atmospheric aspects [3] of data-based models. On the other hand, the geosynchronous meteorological satellite radiation

scanner is a common means of meteorological observation, and its all-weather continuous large-scale observation ability plays an important role in the continuous observation of cloud changes. Because of the narrow observation range of polar orbit satellites, the observation time interval for the same area is large, which makes it impossible to observe the diurnal variation of clouds in the same area and makes usage inconvenient [4]. Therefore, geosynchronous satellites play a useful role in aviation security, disaster prevention, and mitigation. At present, the application level of satellite data is far from what is expected, and massive amounts of satellite observation data are continuously obtained. Therefore, a fast, accurate, and automatic transformation of satellite observation data into usable products becomes very important [5].

The current global cloud classification schemes are not the same, and cloud types are usually divided into four families and ten genera worldwide [6]. However, the classification methods rely on the experience of observers, and the classification results are subjective, which has a definite impact on the classification accuracy. With the continuous progress of satellite observation technology, satellite observation data have been widely used in the field of meteorology [7]. Therefore, the standardization of the global cloud classification scheme is very important. The International Satellite Cloud Climatology Project (ISCCP) proposed a cloud classification standard that is based on satellite and ground-based observations. Cloud types are divided into nine categories by cloud top pressure and cloud optical thickness [8,9]. This classification method reveals the horizontal and vertical characteristics of clouds and provides support for short-term forecasting and climate prediction. Japan and the United States have developed cloud products based on the ISCCP cloud classification scheme for their meteorological satellites. The Japan Aerospace Exploration Agency (JAXA) has developed a cloud classification algorithm based on the ISCCP scheme for its Himawari-8 geosynchronous satellite. Compared with the results of MODIS, it has a high consistency [10,11]. The current CLT product developed by the National Satellite Centre of the China Meteorological Administration for the FY-4A satellite classifies clouds into six categories: water, over cold water, mixed, ice, cirrus, and overlap. The FY-4A satellite has no cloud classification products that are based on the ISCCP cloud classification scheme. Its cloud classification products contain fewer cloud types. Using the ISCCP cloud classification method will greatly enrich the cloud types contained in the current FY-4A cloud classification products. The richer cloud classification products will bring great convenience of application. However, using the current method to develop cloud classification products based on the ISCCP cloud classification scheme for the FY-4A satellite is more complex. The development process needs to consider the influence of different underlying surface types and the characteristics of different observation instruments on cloud classification products, which brings great inconvenience to its use. Therefore, it is necessary to develop a simple cloud classification method.

The current commonly used cloud classification methods mainly include the threshold method, split window method, texture-based method, and mathematical statistics-based method. The split window method and threshold method use reflectivity, brightness temperature, brightness temperature difference values, and underlying surface types as the basis for judging the cloud type [12,13]. These two methods are the most common cloud classification methods. However, because of the complexity of the cloud, these two methods may be at risk of failure in some cases, such as in solar flare areas and high latitude desert areas, where changes in brightness temperature difference can easily lead to misjudgments [14]. For different satellites, these two methods' need to redevelop classification algorithms is highly inconvenient. Texture-based classification methods and mathematical statistics-based methods rely on manpower to find the structures of different cloud types during their development process, which cannot play the role of long-term serial observation data [15,16]. With innovations in technology, the K-means method and SVM method have been introduced into the cloud classification task, and good classification results have been obtained [17,18]. However, the integrity of the cloud is not

fully considered in the classification process. As a continuous whole, the information in the surrounding pixels is very helpful in determining cloud types.

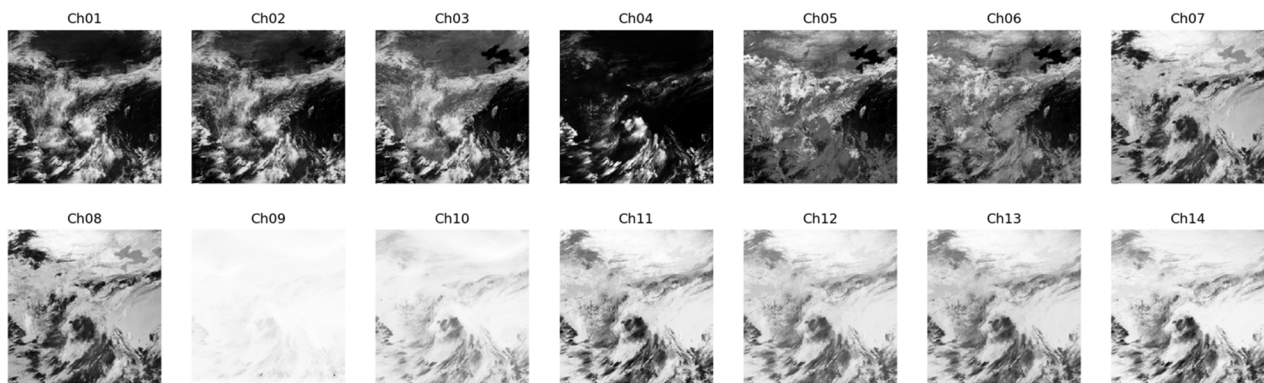
Deep learning technology has achieved excellent results in many fields. It is exceptionally beneficial to use deep learning technology to promote the development of cloud classification methods. The cloud classification methods mainly contain ANN and CNN networks. Taravat A et al. inputted sky camera data into the ANN network for cloud classification tasks and achieved good classification results [19]. Afzali Goroooh et al. used an ANN network to classify satellite images and analyzed the precipitation rates of different types of clouds [20]. Zhang J et al. used CNN networks to identify and classify clouds in images taken from the ground [21]. Among these various methods, because of the limitation of the network itself, the ANN method still ignores the fact that the cloud is a continuous whole, only considering the information under the current pixel. The CNN classification method can only obtain the cloud type of an entire picture and cannot implement pixel-by-pixel cloud classifications, which is inconvenient for fine applications. Zhang Y et al. proposed an improved network based on the U-Net and achieved excellent results for cloud detection of FY-4A observation images [22]. Chai D et al. proposed a deep convolution network based on an encoder–decoder structure for cloud detection for satellite observation images [23]. Drönnner J et al. proposed a CS-CNN network based on image segmentation to implement pixel-by-pixel cloud detection for satellite images [24]. These works achieve good results in cloud detection and have fast detection capabilities. However, these works only solve cloud detection problems and do not solve cloud classification problems. These methods have specific limitations and cannot meet the current application requirements for cloud classification products. Additionally, the models do not pay enough attention to the relationship between the satellite observation channels during the process of cloud detection. The relationship between satellite observation channels is very important for cloud classification tasks.

To solve these problems, we propose a cloud classification method based on the U-Net network. Considering the characteristics of satellite observation data, a channel attention mechanism is used to assign weights to different channels, and a spatial attention mechanism is combined to further improve the accuracy of the classification. The atrous spatial pyramid pooling (ASPP) module is used to enhance the receptive field and strengthen the use of the surrounding pixel information. The long time series FY-4A observations are modeled, and the ISCCP cloud classification scheme is used to classify the FY-4A L1 level data into nine categories, pixel by pixel. The cloud classification products of Himawari-8 and CloudSat are used in the evaluation of the results.

## 2. Dataset

### 2.1. FY-4A Dataset

As a new generation of Chinese geostationary weather satellites, the FY-4A is equipped with an advanced geosynchronous radiation imager (AGRI) and significantly surpasses the previous generation of geostationary satellites in terms of observation time, observation accuracy, and the number of available channels. The AGRI of FY-4A can provide observations at the minute level. During the flood season, FY-4A satellites can scan the Chinese region (REGC) to achieve an observation interval of 5 min, which greatly enhances the ability to observe weather conditions during the flood season [25]. Examples of AGRI observation images are shown in Figure 1.



**Figure 1.** Examples of AGRI all-channel observation images.

The AGRI can observe reflectance and bright temperature information from different wavelengths, and these channels reflect different atmospheric characteristics due to their parameters. The observational parameters for each channel are shown in Table 1.

**Table 1.** FY4A satellite parameters by channel \*.

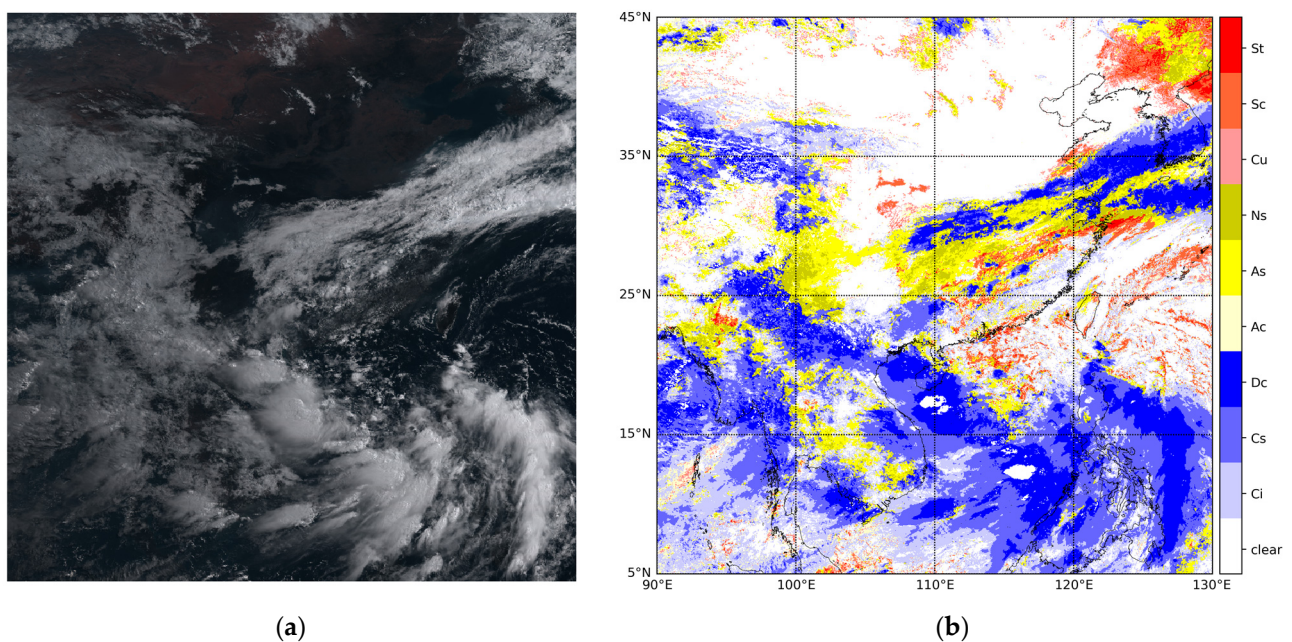
Spectral Coverage	Central Wavelength	Spectral Bandwidth	Spatial Resolution	Main Applications
VIS/NIR	0.47 $\mu\text{m}$	0.45–0.49 $\mu\text{m}$	1 km	Aerosol, visibility
	0.65 $\mu\text{m}$	0.55–0.75 $\mu\text{m}$	0.5 km	Fog, clouds
	0.825 $\mu\text{m}$	0.75–0.90 $\mu\text{m}$	1 km	Aerosol, vegetation
Shortwave IR	1.375 $\mu\text{m}$	1.36–1.39 $\mu\text{m}$	2 km	Cirrus
	1.61 $\mu\text{m}$	1.58–1.64 $\mu\text{m}$	2 km	Cloud, snow
	2.25 $\mu\text{m}$	2.1–2.35 $\mu\text{m}$	2 km	Cloud phase, aerosol, vegetation
Midwave IR	3.75 $\mu\text{m}$	3.5–4.0 $\mu\text{m}$	2 km	Clouds, fire, moisture, snow
	3.75 $\mu\text{m}$	3.5–4.0 $\mu\text{m}$	4 km	Land surface
Water vapor	6.25 $\mu\text{m}$	5.8–6.7 $\mu\text{m}$	4 km	Upper-level WV
	7.1 $\mu\text{m}$	6.9–7.3 $\mu\text{m}$	4 km	Midlevel WV
Longwave IR	8.5 $\mu\text{m}$	8.0–9.0 $\mu\text{m}$	4 km	Volcanic, ash, cloud top, phase
	10.7 $\mu\text{m}$	10.3–11.3 $\mu\text{m}$	4 km	SST, LST
	12.0 $\mu\text{m}$	11.5–12.5 $\mu\text{m}$	4 km	Clouds, low-level WV
	13.5 $\mu\text{m}$	13.2–13.8 $\mu\text{m}$	4 km	Clouds, air temperature

\* Available online: <http://www.sac347.org.cn/nsmc/cn/instrument/AGRI.html> (accessed on 1 March 2022).

## 2.2. Himawari-8 and CloudSat

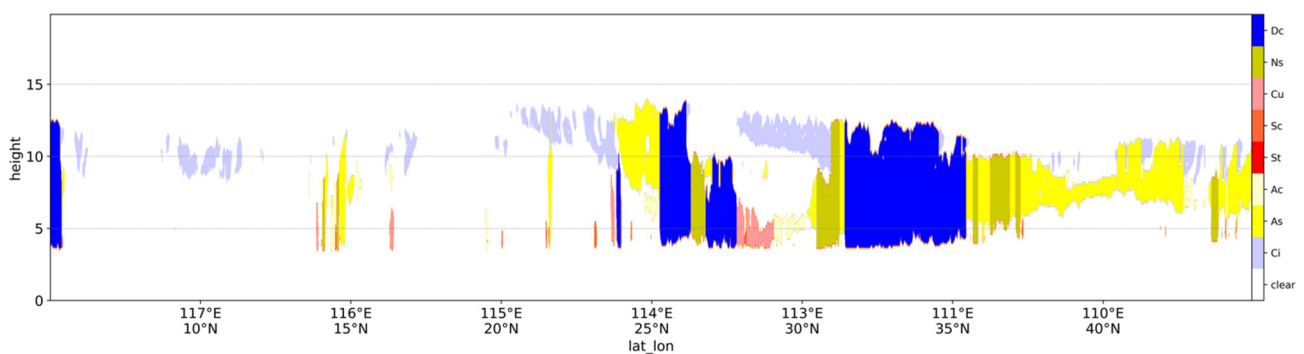
The Himawari-8 satellite is a new generation of geosynchronous satellites developed by the Japan Aerospace Exploration Agency (JAXA), carrying the 16-band Advanced Himawari Imager (AHI) with a spatial resolution of 0.5–2 km [26]. There are differences in the central wavelength distribution compared with the AGRI, which lacks 0.51  $\mu\text{m}$ , 7.3  $\mu\text{m}$ , 9.6  $\mu\text{m}$ , and 11.2  $\mu\text{m}$  wavelengths, but the AHI does not have 1.37  $\mu\text{m}$  wavelengths. JAXA has developed a variety of cloud product data for it. In this paper, the cloud-type product of the Himawari-8 satellite is used as the label for the model training process, and it is also used to test the results of the model [27,28]. The JAXA product is available online: <https://www.eorc.jaxa.jp/ptree/index.html> (accessed on 15 September 2020). We will use Ci (Cirrus), CS (Cirro-Stratus), Dc (Deep-convection), AC (Alto-Cumulus), AS (Alto-Stratus), NS (Nimbo-Stratus), Cu (Cumulus), SC (Strato-Cumulus), and St (Stratus) to indicate the different cloud types in the following. The Himawari-8 satellite observation image and cloud type product are shown in Figure 2.





**Figure 2.** (a) The Himawari-8 satellite observation image; (b) The Himawari-8 satellite cloud-type product.

CloudSat and CALIPSO are equipped with cloud-profiling radar (CPR) and polarization lidar to actively measure the vertical structures inside the clouds. These data greatly contribute to understanding the macrophysical and optical properties of the clouds [29]. The analysis of cloud vertical structures further improves the accuracy of cloud classification tasks [30,31]. Their products have been developed over the years, they are widely recognized for their quality, and they are commonly used in the results inspection process [32]. The cloud type variable visualization for the CloudSat satellite 2B-CLDCLASS is shown in Figure 3.

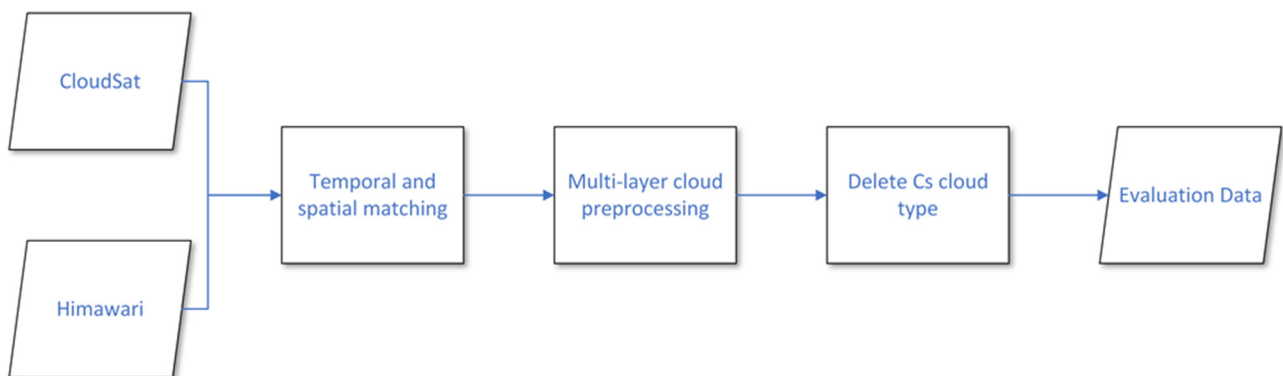


**Figure 3.** Visualization of cloud type variables in the CloudSat satellite 2B-CLDCLASS (Example time 20190708).

The cloud type variable in the 2B-CLDCLASS data from the CloudSat satellite are used in the process of evaluating the classification results, which represents the vertical cloud type structure under the CloudSat orbit [33]. The CloudSat products are available online: <https://www.cloudsat.cira.colostate.edu/data-products/2b-cldclass> (accessed on 17 May 2021). Because CloudSat is a polar-orbiting satellite, the data need to be matched in time and space to the FY-4A satellite when using CloudSat data for results testing. In matching, it is necessary to find the orbit that overlaps with the selected area in this paper, calculate the time when the CloudSat satellite scans into the overlapping area, and find the FY-4A satellite observation that is closest to this time to complete the time matching.

Based on the observation data grid of the FY-4A satellite, the CPR scan point closest to the observation data grid is selected in the CloudSat data, the data are spatially matched, and finally, the data with the completed temporal and spatial matching are used to check the classification results.

Because the CloudSat 2B-CLDCLASS data are the vertical structure of clouds, there will be a situation where multilayer clouds are observed. FY-4A and Himawari-8 satellites are geostationary satellites, they are very difficult to obtain the vertical structure of clouds with, and the obtained data mainly contain information about cloud tops, so we pay more attention to the cloud types in the top layers. In order to use this data for our evaluation process, we needed to preprocess the 2D cloud classification results contained in CloudSat 2B-CLDCLASS to give a single layer of cloud classification results. When there are Dc and NS cloud types in the multilayer clouds, we marked the cloud type at that location as Dc or NS. For other cases of multilayer clouds, we used the higher layer clouds as the cloud type for the current location. CloudSat is missing the CS cloud type compared with the Himawari-8 cloud classification products, and we needed to remove this cloud type from the evaluation process. When the result contained CS cloud type, the point data and CloudSat data are deleted at the same time. The preprocessed cloud-type data are used in the evaluation process. The CloudSat data preprocessing process is shown in Figure 4.

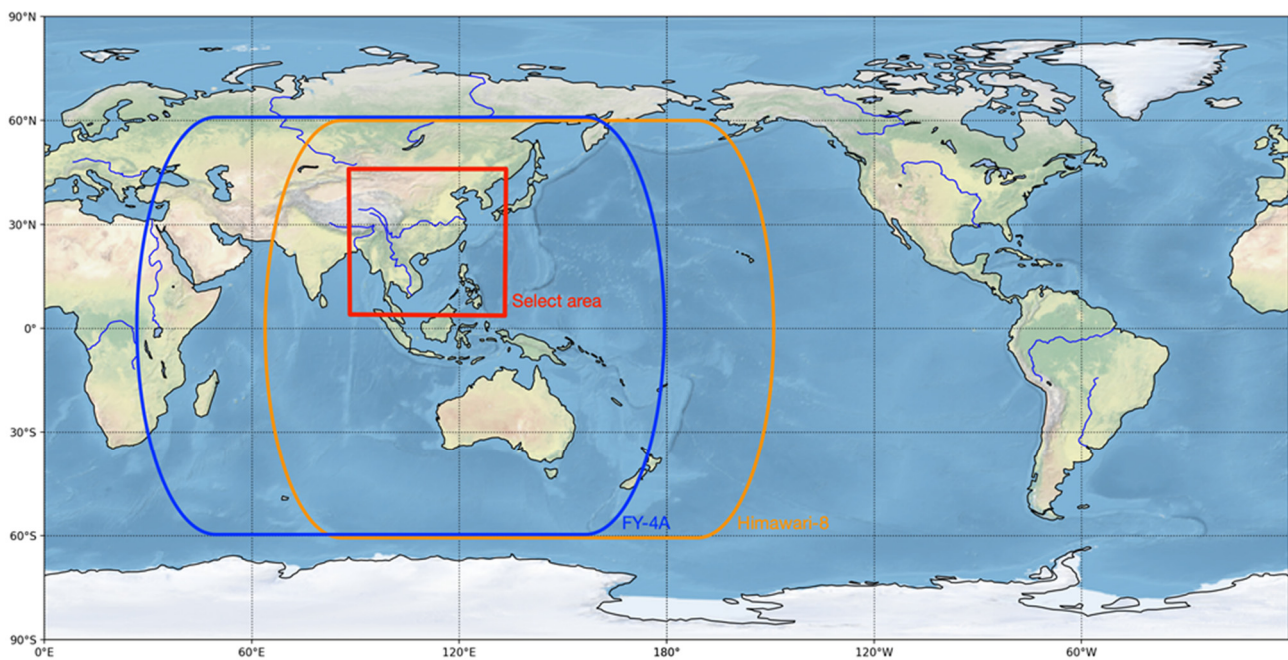


**Figure 4.** The CloudSat data preprocessing process.

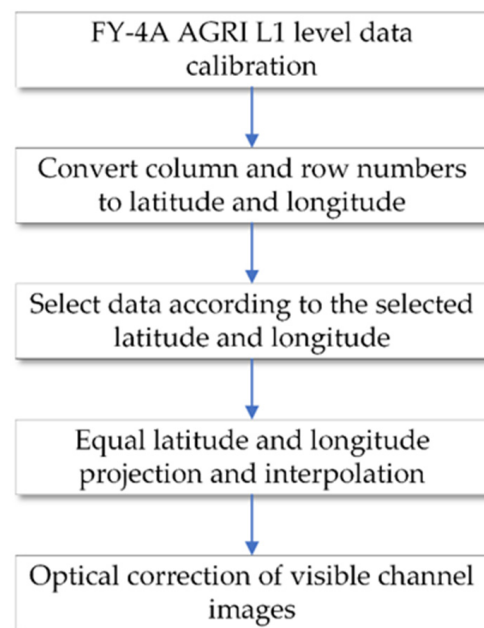
### 2.3. Data Preprocessing

The FY-4A AGRI China region (REGC) 4000 m resolution full disc observations were selected as training data. The raw L1-level data cannot be input directly into the network without calibration, projection conversion, and optical correction steps, so preprocessing is required to obtain the training dataset. The data are calibrated using reflectance calibrations for the visible channel and bright temperature calibrations for the infrared channel. The column and row numbers of the data are converted to latitude and longitude to facilitate the selection of the experimental area. REGC data suffer from observation deficiencies at high latitudes; the range selected is from 5° N to 45° N and 90° E to 130° E. To match the cloud-type data, it is necessary to convert the FY-4A full-disk observations to an equal latitude and longitude projection and then interpolate the converted data to a grid point with a spatial resolution of  $0.05^\circ \times 0.05^\circ$ . The observation areas of the FY-4A and Himawari-8 satellites and the experimentally selected areas are shown in Figure 5.

Finally, the visible channel data are optically corrected. The reflectance observed by the visible channel of the satellite is affected by solar radiation. It needs to be optically corrected using the cosine of the solar zenith angle to ensure that the radiation level in the direct sunlight region is consistent with the dark side. The data preprocessing flow is shown in Figure 6.



**Figure 5.** Observation areas of FY-4A and Himawari-8 satellites and experimentally selected areas.



**Figure 6.** Data preprocessing process.

### 3. Method

This paper proposes a cloud classification algorithm (CLP-CNN) based on the U-Net network. The network incorporates attention mechanisms to improve the capture of interchannel relationships, the use of ASPP modules to expand the network perceptual field, and residual blocks to improve feature extraction. The L1-level observation image from the FY-4A AGRI is simply inputted to classify the current cloud type, taking full advantage of the data-driven approach.

#### 3.1. CLP-CNN

The CLP-CNN network proposed in this paper used the U-Net network as the base structure of the model [34]. This structure allows the network to acquire a pixel-by-



pixel classification capability and is well suited for cloud classification tasks. This paper improves the U-Net network to make it more suitable for cloud classification tasks with satellite-observed images. The CLP-CNN network mainly improves on the U-Net network as follows:

1. To obtain better feature extraction capabilities, the two-dimensional convolution in the U-Net network is replaced with the residual blocks in Res-Net [35,36]. Experiments have shown that deeper convolutional networks give better results, but this improvement brings with it the problem of gradient disappearance and explosion. He et al. 2015 proposed the Res-Net network, which is well solved by the residual structure;
2. To better recognize the information of satellite observation images, the CLP-CNN network adds an attention mechanism to enhance the classification abilities of the network. Xu K et al. proposed an attention mechanism that allows the network to assign higher weights to the desired regions by itself [37]. Hu J et al. 2018 proposed SE-Net, which attempts to improve the finding of the relationships between image channels, a capability well suited for multichannel satellite observation data [38]. Woo S et al. 2018 proposed the convolutional block attention module (CBAM), which enables the network to expand the self-attention along both the channel and spatial dimensions and is of great importance for cloud classification tasks [39];
3. To better integrate features at different levels, information at different scales is merged using the atrous spatial pyramid pooling (ASPP) module [40]. The correlation between the peripheral pixels is exploited;
4. To avoid the information loss caused by the pooling layer during the downsampling process, the CLP-CNN network replaces the pooling layer during the downsampling process with a convolutional layer with a stride of 2, which can lead to finer classification results.

The CLP-CNN network structure is shown in Figure 7.

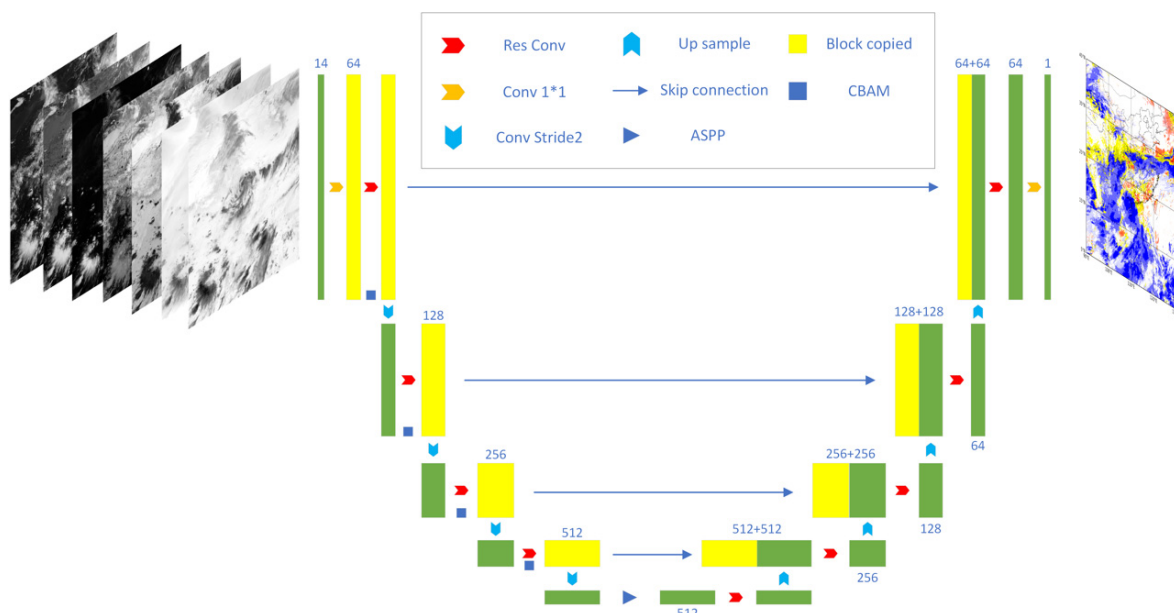
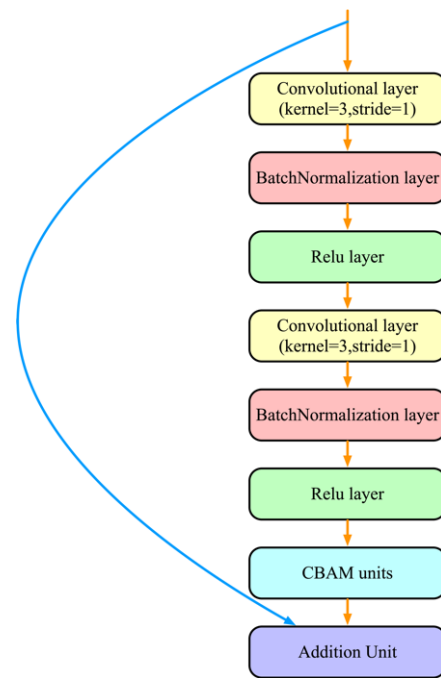


Figure 7. CLP-CNN network structure.

As shown in Figure 7, the model structure is divided into encoder and decoder sections with four downsamples and four upsamples, respectively, corresponding to the left and right sides in Figure 7. In the encoder section, the input data are preprocessed FY-4A observation images containing  $14 \times 800 \times 800$  pixels. The data are adjusted for the number of data channels by means of a  $1 \times 1$  convolutional layer, which avoids causing information loss at the input layer. The encoder section contains a Res-Conv unit in each layer, which

incorporates the residual module and the CBAM module, allowing the network to avoid information loss while increasing the depth of the network. However, the network gains channel and spatial attention. The data from the residual modules and CBAM modules are summed with the original data to complete the feature extraction task. The structure of the Res-Conv unit is shown in Figure 8. During downsampling, the CLP-CNN uses a convolutional layer with a stride of 2 to achieve downsampling, which avoids information loss caused by the pooling layer. The ASPP module was used to extract features at different levels of integration after the last downsampling.



**Figure 8.** Res-Conv unit structure.

In the decoder section, the CLP-CNN uses a skip connection to send the original information to the decoder section and concatenate the original information with the upsampled information. The model will integrate different levels of features. The size of the data is returned to its original size by the Res-Conv module and the upsampling layer, and finally, the number of data channels is changed to 10 layers by a  $1 \times 1$  convolution layer. ( $10 \times 800 \times 800$  pixels). Each pixel point corresponds to a ten-layer probability vector representing the class of the current pixel, and the class with the highest probability for the current pixel will represent the class of that pixel point, ultimately achieving pixel-by-pixel cloud classification.

### 3.2. Train

The train data and test data used in this paper are from FY-4A-observed images and Himawari8 cloud classification products. These data are divided into a training dataset and a test dataset. The training dataset contains FY-4A-observed images and Himawari-8 cloud classification products for 2018 and 2020, where the Himawari-8 cloud classification products will be used as the label during the training process. The testing dataset contains FY-4A-observed images and Himawari-8 cloud classification products for 2019. Meanwhile, we used CloudSat 2B-CLDCLASS data as an additional assessment in the evaluation process. These data are not used in the training and testing process. This is mostly due to the low availability of data from CloudSat satellites and the existence of available data only in 2018 and 2019. All data were preprocessed by the preprocessing method introduced in Sections 2.2 and 2.3, and poor quality and observationally missing data were excluded. The time intervals that are too short have small cloud changes, so the time interval of the



dataset is half an hour. The time range of the dataset was 0–9 UTC. The training dataset contains 10,040 samples, and the testing dataset contains 4540 samples. The evaluation dataset contains 309 samples. The distribution of data are shown in Table 2.

**Table 2.** Distribution of data.

	FY-4A L1	Himawari-8 Cloud Classification Products	CloudSat 2B-CLDCLASS
Training dataset	2018 and 2020	2018 and 2020	Not available
Testing dataset	2019	2019	Not used
Evaluation dataset	2019	2019	2019

As the information observed by different channels of the AGRI has different characteristics, to find the best channel combination plan for cloud classification, this paper designed different channel input combination plans to verify the influence of the number of channels and channel combinations in the input model on the experimental results. The following three plans have been chosen to compare the effects of different channel combinations for input models and are used to determine the channel combinations for input models. The channel combination plans are as follows:

1. The input of all 14 observation channels, including 6 VIS (visible) and 8 IR (infrared) observation channels;
2. The input of all 8 IR observation channels;
3. The input of partial VIS and IR channel data. The plan selects observation channels that have a greater correlation with cloud types, and the selected channels contain 3 VIS channels and 8 IR channels. It contains channels 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, and 14.

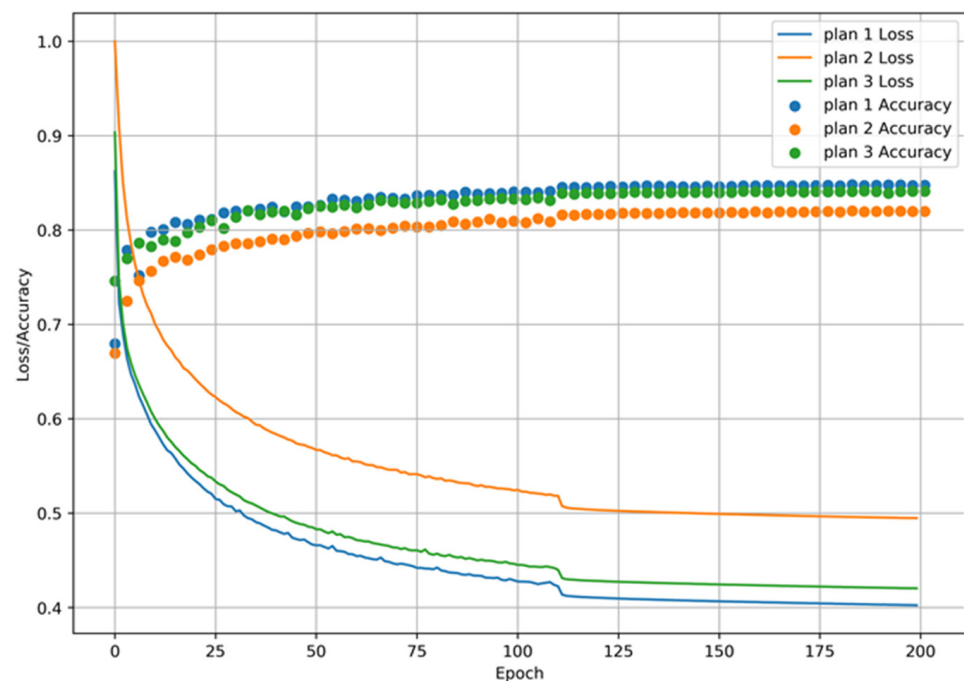
The three-channel combination plans were input into the model for training, and the change in loss between the three plans and the change in accuracy in the test dataset were compared with the channel combination plan to be input into the model for the cloud classification task based on the test results.

Cross entropy was chosen as the loss function of the model, and Adam was chosen as the optimizer of the model in training [41]. The constructed training dataset is input into the model in dimensions of  $\text{batch} \times 14 \times 800 \times 800$  for training. As the number of iterations increases, the variation in the loss and accuracy of the 3-channel combination plan with an increasing number of iterations is shown in Figure 9. A variable learning rate is used in training, and at 110 epochs, the loss and accuracy change significantly as the learning rate changes. Loss is used to describe the bias of the network during the training process. Equation (1) represents the bias for each class in the network. Equation (2) represents the definition of Loss. As shown in Equation (3), accuracy represents the change in accuracy of the testing dataset during the training process.

$$\text{loss}(x, \text{class}) = -\log \left( \frac{\exp(x[\text{class}])}{\sum_j \exp(x[j])} \right) \quad (1)$$

$$\text{Loss} = \frac{\sum_i \text{loss}(x, \text{class}[i])}{i} \quad (2)$$

$$\text{Accuracy} = \frac{\sum_i \text{Hit}[i]}{\text{Total}} \quad (3)$$



**Figure 9.** Variation in the loss and accuracy of training.

The  $i$  in Equations (2) and (3) represent the number of cloud types. The Hit in Equation (3) represents the number of correctly classified points. The Total in Equation (3) represents the total number of points.

#### 4. Result Evaluation

This section evaluates and analyzes the methods proposed in Section 3.2. Several deep learning models are selected and compared with the CLP-CNN network. The effectiveness of different models and different channel combination plans for cloud classifications from FY-4A satellite observation images are evaluated by presenting visual images and evaluation metrics of the classification results output. Quantitative assessments were made using Himawari-8 and CloudSat satellite data; both data types were used to analyze the performance over the seasons.

##### 4.1. Evaluation of CLP-CNN with Different Channel Combination Plans

In this paper, AGRI-observed images of FY-4A from 2019 were chosen to be used as a test set for the model and were used to evaluate the three different channel combination plans presented in Section 3.2. The output of the three plans and the differences in comparing the Himawari-8 CLTYPE products are shown in Figure 10. The outputs of plans 1, 2, and 3 are shown in Figure 10a,c,e, respectively. Figure 10b,d,f separately show the points in the output where the types differ when compared to Himawari-8. These different points are identified in the figures using the classification results of Himawari-8. The blank area is the area where the classification results of both are consistent.

To quantitatively evaluate the classification results of the different channel combination plans, we compared the CLP output, pixel by pixel, with the Himawari-8 product and calculated the confusion matrix. We calculated the POD of the CLP-CNN output compared to the Himawari-8 products of different types using a confusion matrix. The calculated POD is used to evaluate the accuracy of cloud classification tasks under different channel combination plans. The results of the three plans compared with the Himawari-8 product are shown in Table 3. Acc. in Table 3 denotes the accuracy.

According to the statistics in Table 3, Plan 1 achieved the highest accuracy of all the channel combination plans. According to Plan 2, it is clear that the model can only achieve an accuracy of 81.2% in the absence of the VIS channels. In contrast, Plan 3, with the addition of some VIS channels, obtained an accuracy close to that of Plan 1 but is still lower than that of Plan 1. On the other hand, Plan 1 achieves the best POD in all types of clouds. This also shows that the VIS channel has a very positive effect on the cloud classification task and does not cause redundancy in the input information, so this paper uses the full set of observed channel data as input to the CLP-CNN network.

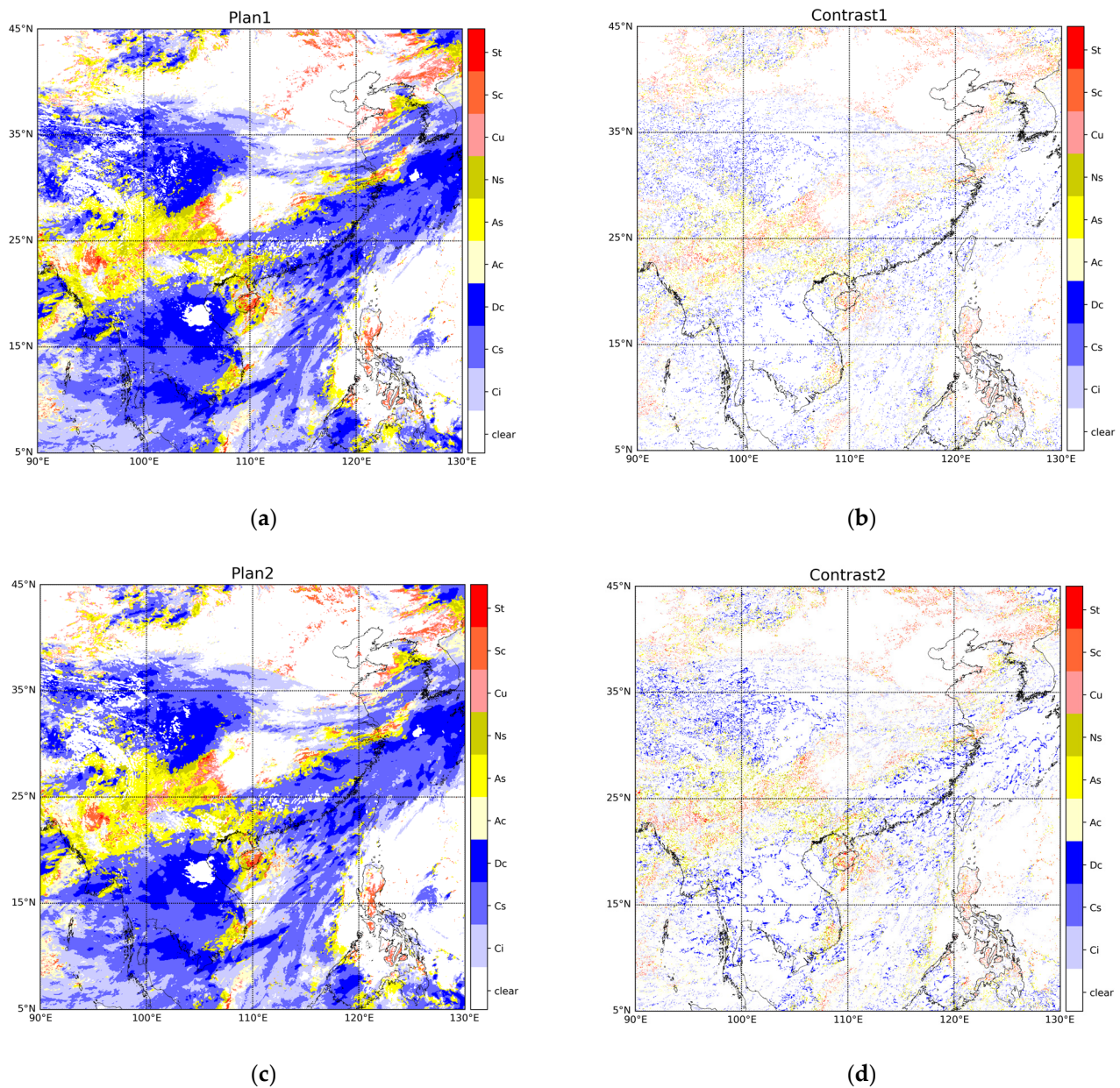
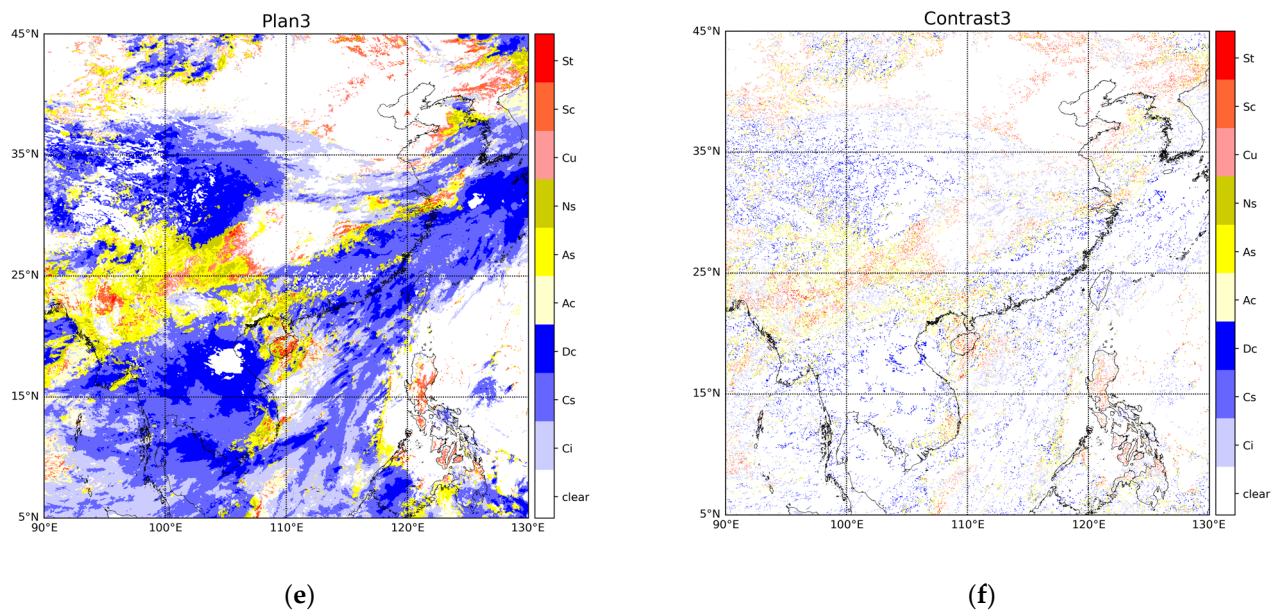


Figure 10. Cont.





**Figure 10.** (a) The output of Plan 1; (b) different points between Plan 1 and Himawari-8; (c) output of Plan 2; (d) different points between Plan 2 and Himawari-8; (e) output of Plan 3; (f) different points between Plan 3 and Himawari-8; (Example time 3 July 2019).

**Table 3.** The results of the three plans compared with the Himawari-8 product.

Type	Plan 1	Plan 2	Plan 3
Clear	0.832	0.801	0.815
Ci	0.614	0.586	0.609
CS	0.834	0.777	0.823
Dc	0.859	0.715	0.836
AC	0.445	0.394	0.447
AS	0.583	0.522	0.566
NS	0.618	0.437	0.593
Cu	0.434	0.399	0.410
SC	0.526	0.415	0.501
St	0.536	0.343	0.512
Acc.	0.844	0.812	0.837

#### 4.2. Comparison of the Results of Different Cloud Classification Models

We used a variety of deep learning models to perform comparative experiments. Improved U-Net-based networks have been widely used in the field of remote sensing and have achieved better results. Therefore, CS-CNN, U-Net++, and U-Net with an attention mechanism are chosen for comparison with the CLP-CNN proposed in this paper. To adapt these networks for comparison to the current task of classifying clouds from satellite observations, the input and output layers of all the models involved in the comparison were modified in this paper to ensure that the networks could input 14 channels of the FY-4A observations and output the correct cloud classification results. All of the comparison models use the same training data, loss function, learning rate, and other parameters. All networks were trained with 200 epochs using the same optimizer, loss function, and learning rate parameters to obtain the classification model. Using the evaluation methods used in Section 4.1, the models in this section were evaluated, and the performance of each of the four models on the cloud classification task was statistically evaluated. The performance of the four models in the cloud classification task is shown in Table 4.

**Table 4.** The POD of the four models compared with the Himawari-8 product.

Type	CS-CNN	U-Net++	U-Net + CBAM	CLP-CNN
Clear	0.799	0.783	0.796	0.832
Ci (Cirrus)	0.561	0.519	0.572	0.614
CS (Cirro-Stratus)	0.802	0.781	0.825	0.834
Dc (Deep-convection)	0.813	0.779	0.836	0.859
AC (Alto-Cumulus)	0.325	0.303	0.352	0.445
AS (Alto-Stratus)	0.573	0.542	0.551	0.583
NS (Nimbo-Stratus)	0.523	0.530	0.598	0.618
Cu (Cumulus)	0.280	0.294	0.319	0.434
SC (Strato-Cumulus)	0.460	0.456	0.458	0.526
St (Stratus)	0.386	0.377	0.354	0.536
Acc.	0.803	0.791	0.818	0.844

The statistical results in Table 4 show that the CLP-CNN achieved the best accuracy rate among all the models. The CLP-CNN achieved a POD of 0.832 in the classification of clear skies, demonstrating its good cloud detection performance compared with the results of 0.799 for CS-CNN, 0.783 for U-Net++, and 0.796 for U-Net+CBAM. The scattered and fragmented nature of these types of clouds makes it difficult to classify them accurately, but CLP-CNN still achieves the best classification results for these types of clouds, demonstrating the excellent cloud classification performance of CLP-CNN networks.

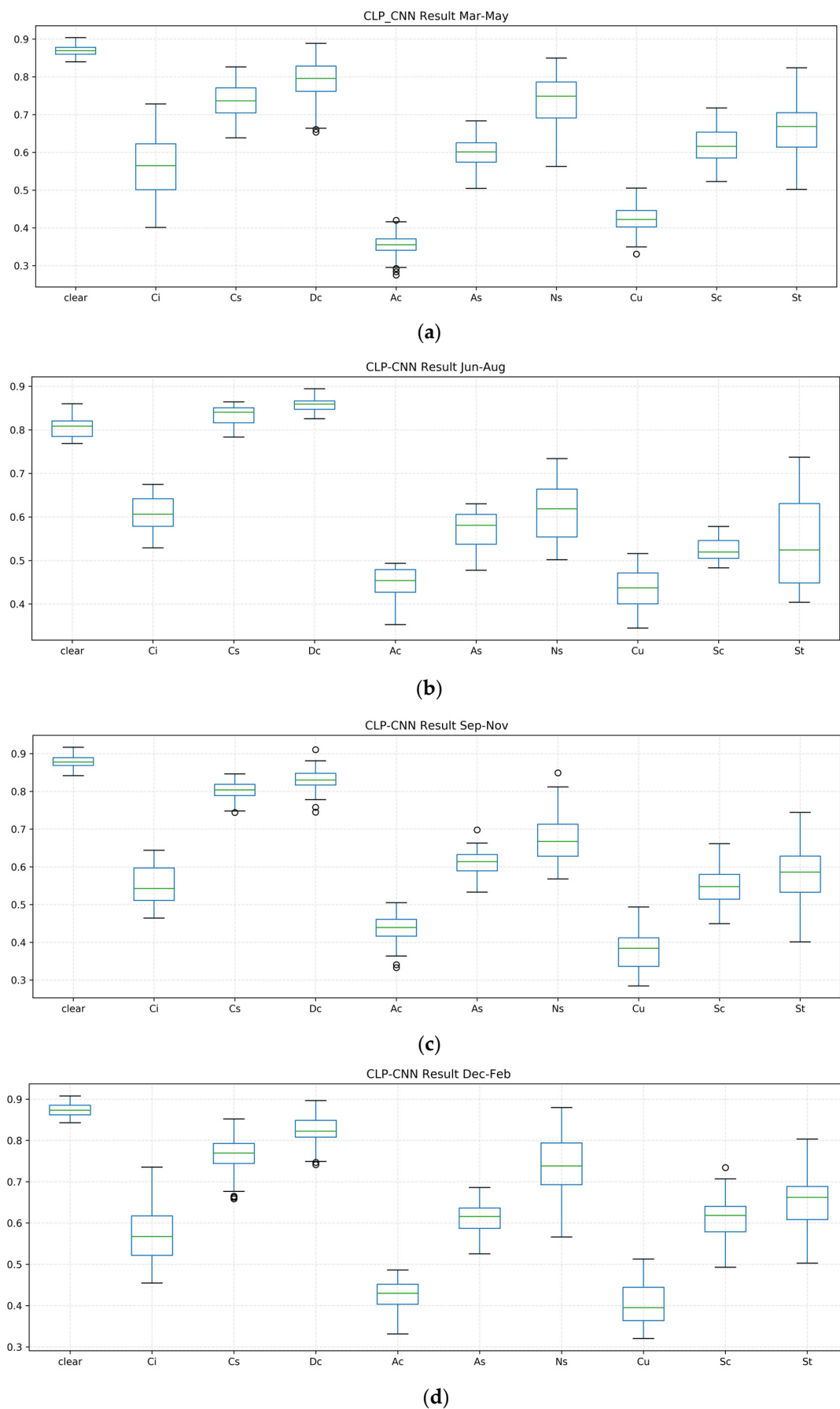
Because of the larger perceptual field resulting from the ASPP structure, the CLP-CNN network makes better use of the information from the surrounding pixels for the cloud classification task and does not incur additional time overhead compared with the other three networks in terms of improved classification accuracy. The CS-CNN does not make use of the attention mechanism to capture the relationship between channels, and its classification results are worse than those of the U-Net network with the CBAM module. The U-Net++ network relies on the pooling layer for its downsampling process compared with the other three networks, and the downsampling process introduces information loss, reducing the network's classification abilities. Therefore, U-Net++ has the least accurate cloud classification rate among the four networks.

#### 4.3. Seasonal Performance Evaluation

In this section, we evaluate the performance of the CLP-CNN network in different seasons. There is a clear seasonal characteristic in the number of clouds as well as the proportion of cloud types in the different seasons in China. This change in distribution will significantly affect the accuracy of the CLP-CNN network in classifications across the seasons. Considering the proportion of different types of clouds in different seasons, we evaluate cloud classification models according to March–May, June–August, September–November, and December–February 2019. The classification results of these seasons are evaluated using POD as an evaluation criterion. The results of the classification evaluation for the different seasons are shown in Figure 11.

Based on the statistical results of the seasonal evaluation, the different cloud types are strongly influenced by their distribution characteristics. The results of the middle-level and high-level cloud classifications showed high robustness throughout the year. However, there were some fluctuations in the correct rate of low-level and middle-level clouds over the seasons. For example, the St cloud type has a high POD score in March–May, but in June–August, the St distribution decreases significantly compared with March–May, its POD score decreases, and its accuracy fluctuates from time to time. In the December–February period, the St distribution share improved by 54%. Therefore, there was a further increase in POD. The statistics on the percentage of cloud types also showed that the percentage of different types of clouds showed a significant change in different seasons. Therefore, the seasons had some influence on the accuracy of the classification results of the CLP-CNN network.





**Figure 11.** Results of the classification evaluation for the different seasons.

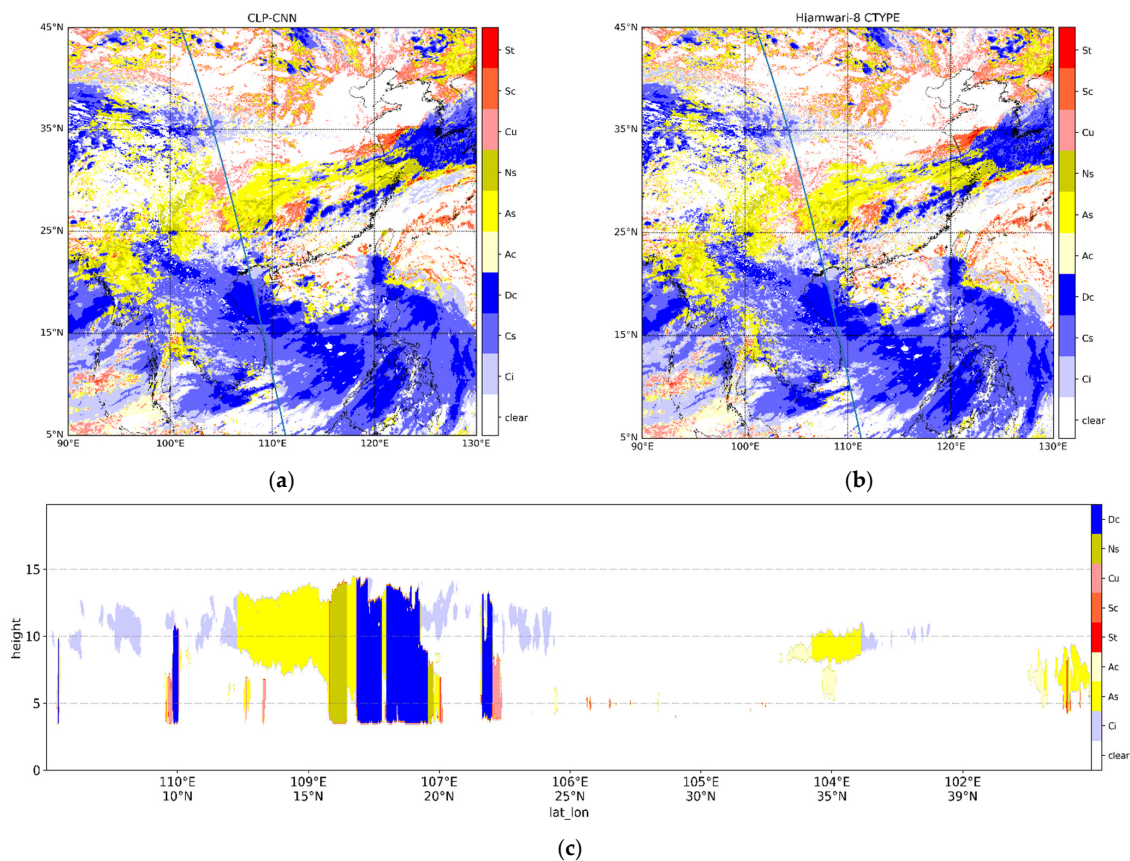
The Ci, AC, and Cu cloud types are thin in optical thickness compared with the other cloud types in the ISCCP cloud classification standard. The distribution of the three types of clouds is also more dispersed compared with the other types of clouds, and they present a greater challenge to the cloud classification model. All three types of clouds do not give similar results relative to other cloud types of similar cloud top heights. In contrast, the Clear Sky, St, NS, and Dc types have a larger proportion of cloud types, while they are generally located in the center of the cloud mass and are more continuous in their distribution, thus achieving a higher percentage of accuracy. The Ci, AC, and Cu types of clouds are generally distributed at the edge of the cloud cluster, the distribution is discontinuous and fragmented, and their proportion is lower than other types of clouds. This leads to their low proportion in the training samples, and there is a serious problem of an uneven distribution of samples. In addition, parallax errors of geostationary satellites also affect the classification accuracy of these thin clouds. These reasons have caused the cloud classification accuracy to be below expectations of the more effective evaluation accuracy of the CLP-CNN network output. The errors are not only from the FY-4A satellite observations but also from the Himawari-8 satellite. In order to better verify the accuracy of cloud classification results of the CLP-CNN network, we used CloudSat data for evaluation in Section 4.4.

#### 4.4. CloudSat

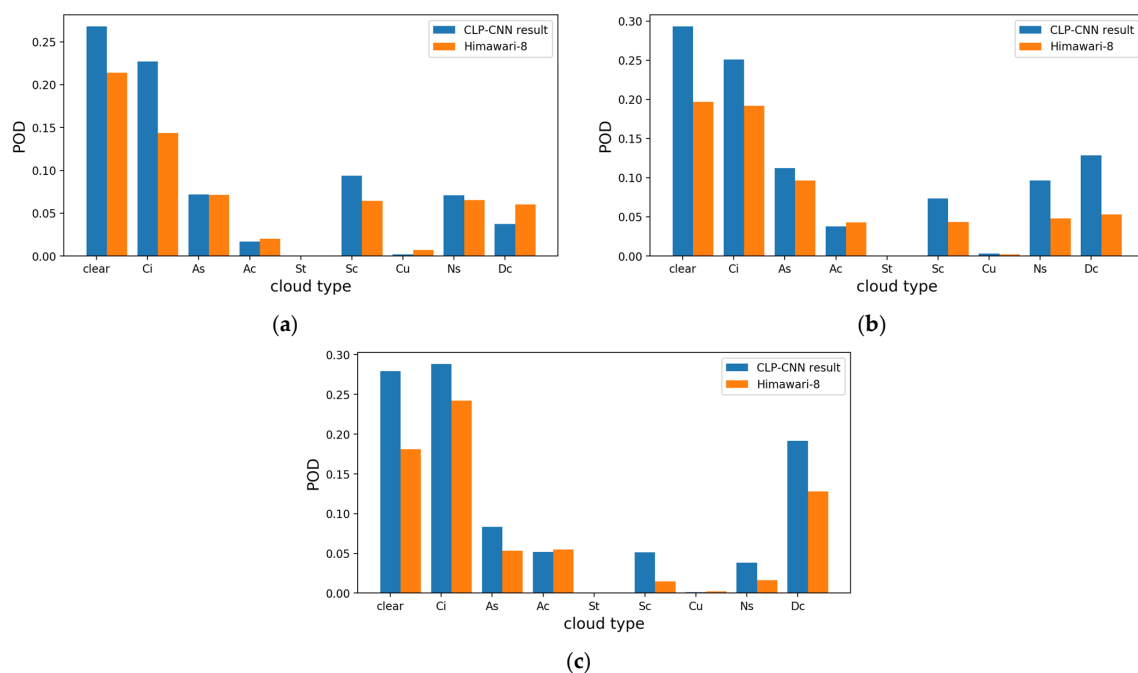
In this section, 2B-CLDCLASS data from the Cloudsat satellite were used to evaluate the results. With the CPR radar on board, the CloudSat satellite can accurately determine the type of clouds in the cloud layer below. We use the CloudSat classification results to better evaluate the accuracy of the CLP-CNN cloud classification results. The output of the CLP-CNN network and the Himawari-8 product were matched to the CloudSat data using the time–space matching method described in Section 2.2. The confusion matrix of the matched data is calculated to obtain the evaluation results.

A comparison of the output of the CLP-CNN and the Himawari-8 product using January–July 2019 CloudSat data show a high similarity between the output of the CLP-CNN network and the Himawari-8 product. Using the classification results from the CloudSat satellite as true values, the accuracy of the CLP-CNN output reached 0.523, which is better than the 0.517 of the Himawari-8 products. Figure 12 shows one of the examples. The comparison in the figure shows that both the CLP-CNN and Himawari-8 products recognize AS and NS at approximately  $15^{\circ}$  N as Cs. A similar situation occurs in the region at approximately  $35^{\circ}$  N. We believe this is due to the inability of synoptic satellites to detect the detailed structure of cloud insides, which can lead the network to misclassify cloud types when the cloud top height and cloud optical thickness of the classified clouds are similar to those of deep convective clouds. In other regions, both products accurately classify the current cloud type.

To more clearly evaluate the accuracy of the CLP-CNN output cloud-classification results, we selected different points from the CLP-CNN output results and the Himawari-6 products. We conducted a seasonal test using CloudSat 2B-CLDCLASS data from January to July 2019. The selected points with different classifications were matched with CPR scanning points closest to the point and compared with CloudSat. The comparison results were evaluated by POD. The statistical results of the evaluations are shown in Figure 13.



**Figure 12.** Comparison of the CLP-CNN output and Himawari-8 product with the CloudSat 2B-CLDCLASS product. The blue line in (a,b) represents the satellite orbit of CloudSat; (a) CLP-CNN output results; (b) Himawari-8 product; (c) CloudSat 2B-CLDCLASS product; (example time 20190701).



**Figure 13.** (a) Comparison of CLP-CNN results and Himawari-8 products with CloudSat from January to February; (b) comparison of CLP-CNN results and Himawari-8 products with CloudSat from March to May; (c) comparison of CLP-CNN results and Himawari-8 products with CloudSat from June to July.

By comparing the output results of the CLP-CNN and Himawari-8 products shown in Figure 13 with CloudSat, it can be found that the output results of the CLP-CNN network are worse than those of the Himawari-8 products only in the classification of AC-type clouds. In the regions with different classifications, the cloud classification results of CLP-CNN are closer to those of CloudSat. The reason for this problem is also the difference between the satellites and the feature extraction abilities of the CLP-CNN network for long-term sequence data. Therefore, the CLP-CNN network has a good classification effect on the observation data of geosynchronous satellites.

## 5. Conclusions

In this paper, the CLP-CNN network is proposed, a CNN network that can be used for cloud classifications of satellite observation images. This model combines an attention mechanism, a residual network, and an ASPP module to improve the classification abilities of the network. After training, the CLP-CNN achieved an 84.4% accuracy compared with the Himawari-8 cloud classification products; when evaluated with CloudSat 2B-CLDCLASS, the CLP-CNN output was superior to the Himawari-8 products for various cloud types. Compared with many excellent classification models in the field of remote sensing, the CLP-CNN network achieves the best results. The model only needs 0.9 s to perform the cloud classification task on the FY-4A AGRI observation image with  $800 \times 800$  pixels without manual intervention and professional knowledge. However, the CLP-CNN network has some shortcomings. Because of the influence of CloudSat's running orbit and running time, we cannot introduce more accurate cloud classification data into the training process. We plan to further improve our label dataset in the future to make it more accurate and strengthen the model to address uneven distributions of samples. At the same time, significant differences in seasons also had an impact on the classification ability of the CLP-CNN network. We will try to introduce seasonal characteristics into the model training process to improve the classification performance of the network.

**Author Contributions:** Conceptualization, Y.J. and W.C.; methodology, Y.J. and W.C.; software, Y.J.; validation, Y.J., W.C. and F.G.; formal analysis, Y.J.; investigation, S.W.; resources, W.C. and F.G.; data curation, Y.J. and S.W.; writing—original draft preparation, Y.J.; writing—review and editing, Y.J., W.C. and J.L.; visualization, Y.J. and W.C.; supervision, S.Z. and C.L.; project administration, S.Z.; funding acquisition, F.G. and J.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the National Key R&D Program of China (Grant numbers 2021YFC3101505) and the National Natural Science Foundation of China (Grant No. 41830964), and the Shandong Province's "Taishan" Scientist Project (ts201712017), and Qingdao "Creative and Initiative" frontier Scientist Program (19-3-2-7-zhc).

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank the China National Satellite Meteorological Center, the Himawari-8 data website, and the CloudSat data website, which are freely accessible to the public. The research product of the cloud type that was used in this paper was supplied by the P-Tree System, Japan Aerospace Exploration Agency (JAXA).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Baker, M. Cloud microphysics and climate. *Science* **1997**, *276*, 1072–1078. [\[CrossRef\]](#)
2. Min, M.; Wang, P.; Campbell, J.R.; Zong, X.; Li, Y. Midlatitude cirrus cloud radiative forcing over China. *J. Geophys. Res. Atmos.* **2010**, *115*, 1408–1429. [\[CrossRef\]](#)
3. Liu, Y.; Xia, J.; Shi, C.-X.; Hong, Y. An improved cloud classification algorithm for China's FY-2C multi-channel images using artificial neural network. *Sensors* **2009**, *9*, 5558–5579. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Stubenrauch, C.J.; Rossow, W.B.; Kinne, S.; Ackerman, S.; Cesana, G.; Chepfer, H.; Di Girolamo, L.; Getzewich, B.; Guignard, A.; Heidinger, A. Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX radiation panel. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1031–1049. [\[CrossRef\]](#)

5. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204. [\[CrossRef\]](#)
6. Huertas-Tato, J.; Rodríguez-Benítez, F.J.; Arbizu-Barrena, C.; Aler-Mur, R.; Galvan-Leon, I.; Pozo-Vázquez, D. Automatic Cloud-Type Classification Based on the Combined Use of a Sky Camera and a Ceilometer. *J. Geophys. Res. Atmos.* **2017**, *122*, 11045–11061. [\[CrossRef\]](#)
7. Min, M.; Bai, C.; Guo, J.; Sun, F.; Liu, C.; Wang, F.; Xu, H.; Tang, S.; Li, B.; Di, D.; et al. Estimating Summertime Precipitation from Himawari-8 and Global Forecast System Based on Machine Learning. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2557–2570. [\[CrossRef\]](#)
8. Rossow, W.B.; Schiffer, R.A. ISCCP cloud data products. *Bull. Am. Meteorol. Soc.* **1991**, *72*, 2–20. [\[CrossRef\]](#)
9. Hahn, C.J.; Rossow, W.B.; Warren, S.G. ISCCP cloud properties associated with standard cloud types identified in individual surface observations. *J. Clim.* **2001**, *14*, 11–28. [\[CrossRef\]](#)
10. Mouri, K.; Izumi, T.; Suzue, H.; Yoshida, R. Algorithm Theoretical Basis Document of cloud type/phase product. *Meteorol. Satell. Cent. Tech. Note* **2016**, *61*, 19–31.
11. Suzue, H.; Imai, T.; Mouri, K. High-resolution cloud analysis information derived from Himawari-8 data. *Meteorol. Satell. Cent. Tech. Note* **2016**, *61*, 43–51.
12. Ackerman, S.A.; Strabala, K.I.; Menzel, W.P.; Frey, R.A.; Moeller, C.C.; Gumley, L.E. Discriminating clear sky from clouds with MODIS. *J. Geophys. Res. Atmos.* **1998**, *103*, 32141–32157. [\[CrossRef\]](#)
13. Purbantoro, B.; Aminuddin, J.; Manago, N.; Toyoshima, K.; Lagrosas, N.; Sumantyo, J.T.S.; Kuze, H. Comparison of Cloud Type Classification with Split Window Algorithm Based on Different Infrared Band Combinations of Himawari-8 Satellite. *Adv. Remote Sens.* **2018**, *7*, 218–234. [\[CrossRef\]](#)
14. Poulsen, C.; Egede, U.; Robbins, D.; Sandeford, B.; Tazi, K.; Zhu, T. Evaluation and comparison of a machine learning cloud identification algorithm for the SLSTR in polar regions. *Remote Sens. Environ.* **2020**, *248*, 111999. [\[CrossRef\]](#)
15. Zhang, C.; Yu, F.; Wang, C.; Yang, J. Three-dimensional extension of the unit-feature spatial classification method for cloud type. *Adv. Atmos. Sci.* **2011**, *28*, 601–611. [\[CrossRef\]](#)
16. Gao, T.; Zhao, S.; Chen, F.; Sun, X.; Liu, L. Cloud Classification Based on Structure Features of Infrared Images. *J. Atmos. Ocean. Technol.* **2011**, *28*, 410–417. [\[CrossRef\]](#)
17. Heinle, A.; Macke, A.; Srivastav, A. Automatic cloud classification of whole sky images. *Atmos. Meas. Tech.* **2010**, *3*, 557–567. [\[CrossRef\]](#)
18. Gómez-Chova, L.; Camps-Valls, G.; Bruzzone, L.; Calpe-Maravilla, J. Mean map kernel methods for semisupervised cloud classification. *IEEE Trans. Geosci. Remote Sens.* **2009**, *48*, 207–220. [\[CrossRef\]](#)
19. Taravat, A.; Del Frate, F.; Cornaro, C.; Vergari, S. Neural networks and support vector machine algorithms for automatic cloud classification of whole-sky ground-based images. *IEEE Geosci. Remote Sens. Lett.* **2014**, *12*, 666–670. [\[CrossRef\]](#)
20. Afzali Goroooh, V.; Kalia, S.; Nguyen, P.; Hsu, K.-I.; Sorooshian, S.; Ganguly, S.; Nemani, R. Deep Neural Network Cloud-Type Classification (DeepCTC) Model and Its Application in Evaluating PERSIANN-CCS. *Remote Sens.* **2020**, *12*, 316. [\[CrossRef\]](#)
21. Zhang, J.; Liu, P.; Zhang, F.; Song, Q. CloudNet: Ground-based cloud classification with deep convolutional neural network. *Geophys. Res. Lett.* **2018**, *45*, 8665–8672. [\[CrossRef\]](#)
22. Zhang, Y.; Cai, P.; Tao, R.; Wang, J.; Tian, W. Cloud Detection for Remote Sensing Images Using Improved U-Net. *Bull. Surv. Mapp.* **2020**, *3*, 17–20. [\[CrossRef\]](#)
23. Chai, D.; Newsam, S.; Zhang, H.K.; Qiu, Y.; Huang, J. Cloud and cloud shadow detection in Landsat imagery based on deep convolutional neural networks. *Remote Sens. Environ.* **2019**, *225*, 307–316. [\[CrossRef\]](#)
24. Drönner, J.; Korfhage, N.; Egli, S.; Mühling, M.; Thies, B.; Bendix, J.; Freisleben, B.; Seeger, B. Fast Cloud Segmentation Using Convolutional Neural Networks. *Remote Sens.* **2018**, *10*, 1782. [\[CrossRef\]](#)
25. Guo, Q.; Lu, F.; Wei, C.; Zhang, Z.; Yang, J. Introducing the New Generation of Chinese Geostationary Weather Satellites, Fengyun-4. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 1637–1658. [\[CrossRef\]](#)
26. Bessho, K.; Date, K.; Hayashi, M.; Ikeda, A.; Imai, T.; Inoue, H.; Kumagai, Y.; Miyakawa, T.; Murata, H.; Ohno, T. An introduction to Himawari-8/9—Japan’s new-generation geostationary meteorological satellites. *J. Meteorol. Soc. Jpn. Ser. II* **2016**, *94*, 151–183. [\[CrossRef\]](#)
27. Ishida, H.; Nakajima, T.Y. Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager. *J. Geophys. Res.* **2009**, *114*, 141–157. [\[CrossRef\]](#)
28. Letu, H.; Nagao, T.M.; Nakajima, T.Y.; Riedi, J.; Ishimoto, H.; Baran, A.J.; Shang, H.; Sekiguchi, M.; Kikuchi, M. Ice Cloud Properties from Himawari-8/AHI Next-Generation Geostationary Satellite: Capability of the AHI to Monitor the DC Cloud Generation Process. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 3229–3239. [\[CrossRef\]](#)
29. Min, M.; Wang, P.; Campbell, J.R.; Zong, X.; Xia, J. Cirrus cloud macrophysical and optical properties over North China from CALIOP measurements. *Adv. Atmos. Sci.* **2011**, *28*, 653–664. [\[CrossRef\]](#)
30. Sassen, K.; Wang, Z.; Liu, D. Global distribution of cirrus clouds from CloudSat/Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) measurements. *J. Geophys. Res.* **2008**, *113*, D8. [\[CrossRef\]](#)
31. Powell, K.A.; Hu, Y.; Omar, A.; Vaughan, M.A.; Winker, D.M.; Liu, Z.; Hunt, W.H.; Young, S.A. Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms. *J. Atmos. Ocean. Technol.* **2009**, *26*, 2310–2323. [\[CrossRef\]](#)



32. Vane, D.; Stephens, G.L. The CloudSat Mission and the A-Train: A Revolutionary Approach to Observing Earth's Atmosphere. In Proceedings of the 2008 IEEE Aerospace Conference, Big Sky, MT, USA, 1–8 March 2008; pp. 1–5. [[CrossRef](#)]
33. Wang, Z.; Sassen, K. Level 2 cloud scenario classification product process description and interface control document. *CloudSat Proj. NASA Earth Syst. Sci. Pathfind. Mission* **2007**, *5*, 50.
34. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Lecture Notes in Computer Science; Springer Science + Business Media: Berlin, Germany, 2015; pp. 234–241.
35. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
36. He, K.; Zhang, X.; Ren, S.; Sun, J. Identity mappings in deep residual networks. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; pp. 630–645.
37. Xu, K.; Ba, J.; Kiros, R.; Cho, K.; Courville, A.; Salakhudinov, R.; Zemel, R.; Bengio, Y. Show, attend and tell: Neural image caption generation with visual attention. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 2048–2057.
38. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7141.
39. Woo, S.H.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In Proceedings of the Computer Vision—ECCV 2018, PT VII, Munich, Germany, 8–14 September 2018; pp. 3–19.
40. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.
41. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.