

Classification of Ice Crystal Images from Airborne Cloud Particle Imager Probe (CPI) Using Convolutional Neural Networks (CNN)

Ye-Feng Xu¹, Rui-Li Jiao¹, Min-Song Huang^{2,3,4*}

¹ School of Information and Communication Engineering, Beijing Information Science and Technology University, Beijing 100101, China

{yefeng.xu, jiaoruili}@bistu.edu.cn

² Key Laboratory of Transportation Meteorology of China Meteorological Administration, Nanjing Joint Institute for Atmospheric Sciences, Nanjing 210041, China

³ Key Laboratory of Cloud-Precipitation Physics and Weather Modification (CPML), CMA, Beijing 100081, China

⁴ Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

mission@mail.iap.ac.cn

Received 27 March 2024; Revised 18 April 2024; Accepted 23 May 2024

Abstract. Ice crystals in clouds have various shapes, which play a crucial role in understanding the development of precipitation, climate change and remote sensing retrievals. The copious ice crystal images collected by the airborne cloud particle imager probe (CPI) following each research flight impede efficient human identification, prompting the necessity for an automated, high-precision algorithm to classify ice crystal habits. Traditional automatic classification methods require manual feature extraction for a good performance, which affects their generalization ability. Instead, the recently perfected machine learning method -- convolutional neural network (CNN) holds promise in addressing this issue. In this paper, the ice crystal images observed by CPI are used to set up an ice crystal dataset, which consists of eleven shapes containing 5,342 images. Additionally, a method to identify ice particle shape based on CNN is presented. The small 3×3 convolutional kernels are used to construct a 30-layer CNN model to achieve automatic habit classification of ice crystal particle shapes. The CNN model is compared with traditional machine learning models (SVM, BP) using the created dataset. The CNN model achieved the highest F1 score for each category and an accuracy of 95.45%. Experimental results show that ice crystal classification using CNN is an effective and feasible method, surpassing traditional classification methods that require manual feature extraction. This research provides a reference value for cloud microphysics research.

Keywords: ice crystal habits, Cloud Particle Imager (CPI), convolutional neural networks (CNN), cloud microphysics research, deep learning

1 Introduction

Ice clouds have an important influence on atmospheric radiation and energy transport. The radiative characteristics of ice crystal particles are mostly determined by their size and habit [1]. The radiative effects of cirrus clouds have been calculated with the help of the distribution of ice cloud particle habits [2]. The net radiative effect of cirrus clouds is influenced by the size distribution and ice crystal habit inside the cloud as well as the cloud's vertical and horizontal dimensions [3]. Additionally, the morphology and size of ice crystal particles allow for the calculation of cloud microphysical characteristics in different ice clouds, such as particle spectral distribution, ice-water content, liquid-water content, precipitation, effective particle radius, and extinction coefficient [4]. And the parameters of different ice crystal shapes play a crucial role in determining cloud parameterization schemes in numerical models and cloud radiation characteristics in atmospheric remote sensing [4]. When estimating the cloud's overall mass and optical thickness, the morphology of the ice crystal particles is crucial, especially when evaluating how individual crystals interact with electromagnetic radiation [5]. Additionally, distinct ice crystal forms correspond to distinct microphysical and radiative properties as well as having distinct single scattering properties [6, 7]. Studying the form habits of ice crystal particles within clouds is crucial as a result.

* Corresponding Author

Aircraft observation is one of the important methods for measuring the size and morphology of ice crystal particles in clouds, especially with the airborne imaging probes. However, the airborne imaging probes can only capture the two-dimensional projection image and size of ice crystal particles, without the ability to identify their shapes. While ice particle shapes can be identified by the human eye, it is very time-consuming and subject to human subjectivity, because the imaging probe can usually record tens of thousands of ice particles after each flight. Therefore, it is necessary to develop an automatic ice crystal habit recognition technology.

In the study of automatic identification methods for ice crystal particle shapes, Korolev and Sussman [8] employed dimensionless ratio relationships and an inverse approach to classify ice crystal particles into four shapes, utilizing the obtained feature parameters from particle images. Lawson et al. [9] presented a method to classify the particle shapes by using simple geometric dimensions of the particles in a combinatorial operation. Praz et al. [10] utilized polynomial logistic regression with geometric and textural descriptions to categorize six ice crystal shapes. In China, Huang and Lei [11] presented an enhanced Holroyd method for identifying cloud particle habits. Dong et al. [12] utilized a back propagation (BP) neural network optimized by simulated annealing algorithm to distinguish the shape of cloud particle images. Nevertheless, traditional classification methods often require a set of extracted features rather than directly using ice crystal images as input. For complex structures like ice crystal particles, it may be difficult to design effective feature representations, which will affect the generalization performance of the algorithm.

Convolutional neural network (CNN), in particular, are deep learning algorithms that have attracted a lot of interest and applications recently due to their capacity to train and categorize image recognition without the need for the pretreatment phase of picture feature extraction. Currently, this technique has been applied to classification of ice crystal particle images. Touloupas et al. [13] employed a CNN to classify ice crystals sampled by holographic imager into three categories with an accuracy of over 96%. Liao et al. [14] designed Hy-INet, a CNN embedded with a hypergraph convolution module, achieving a remarkable accuracy of 98.08% in a ten-class classification of cloud particles using CPI data. Jiao et al. [15] constructed a three-layer CNN model to identify supercooled droplet images captured by CPI with an accuracy of 99.5%. Huo et al. [16] used lightweight convolution modules to build CNN model that classified nine cloud particle shapes sampled by two-dimensional stereo probe detector (2D-S) with 96% accuracy. Xiao et al. [17] and Przybylo et al. [18] employed transfer learning technique to classify ice crystal images sampled by CPI. Wu et al. [19] also utilized transfer learning approach to classify particle images sampled by the Cloud Imaging Probe (CIP). The classification accuracies were all higher than 96%. All these preceding studies demonstrate that CNNs necessitate minimal image preprocessing and eliminate the need for manually extracted features. They autonomously extract features from the data, enabling more dependable and robust predictions directly from ice crystal particle images. In this study, we first established an ice crystal database comprising 11 distinct categories, totaling 5342 images. The images are acquired by the airborne CPI over continental regions of China and North America, as well as the Pacific and Atlantic Oceans. And then we constructed a CNN-based classification model, which was trained and optimized using our constructed ice crystal particle shape database. With simple image preprocessing, reasonable network architecture design and parameter tuning, we obtained an accuracy of 95.45%.

The rest of this paper is organized as follows. The database and probing device's working principle is explained in Section 2. Section 3 introduces the CNN method for automatic classification of ice crystal particles for CPI probes in detail, including image preprocessing, experimental platform, and a description of the training process. The experimental comparison mode and the process for assessing model performance are covered in Section 4. Section 5 provides the results of the experimental comparisons. Section 6 summarizes the work of this paper.

2 Instrument and Dataset

Observing the size and habit distribution of single ice crystal within cloud is difficult. However, the size and habits of cloud particles in clouds can be measured by airborne particle imaging instruments, such as the Stratton Park Engineering Company's (SPEC) CPI. The particle detection system of CPI consists of two continuous wave laser diodes and their corresponding detectors crossing vertically to form a 2.4 mm wide, 0.5 mm thick strip laser beam. The imaging pulsed laser emits a pulsed beam of light in response to a particle passing at the junction of the two laser beams; the particle image is then projected onto the CCD. Furthermore, CPI can detect sizes from 2.3 to 2300 μm and is quantized at 8 bits by the CCD camera. Its maximum transmission rate is 400 frames per second, allowing for fast imaging of smaller particles. Additionally, CPI is resistant to low pressure and low

temperature. More information on CPI probe can be found in Glienke and Mei [20].

Using images of ice crystal particles observed by CPI over continental regions of China and North America, as well as the Pacific and Atlantic Oceans, we created a geographically diverse image dataset containing 11 ice crystal categories. Data from continental China were sampled by the Beijing Weather Modification Office's and the Hebei Provincial Weather Modification Office's King-air research plane in East and North China, respectively. Data from North America and the Pacific and Atlantic Oceans were collected from five field observation projects: ATTREX, IDAEAS-4, POSIDON, PREDICT and SEAC4RS. Table 1 provides specific information about these projects. A dataset containing a sufficient number of precisely labeled ice crystal particles is needed to create an automatic shape identification system for ice crystal particles. In this study, an ice crystal habit dataset is constructed manually. Table 2 displays the specifics of the ice crystal shape dataset. 5342 representative images of CPI ice crystal particles are included in the dataset; only 200 images belong to the dendritic category. To ensure a balanced number of samples for each class during model training, 200 particle images of each of the 11 categories of ice crystal particles were randomly selected from the dataset for model training and testing.

Table 1. Details of CPI data acquisition for the outfield observation projects

Campaign	Location	Date	Aircraft	Reference	Major ice crystals found
ATTREX	Western Pacific	Jan - Mar 2014	Global Hawk	Jensen et al. [21]	Pla.Agg., Ros, Bud, Col, Agg
IDAEAS-4	North Central U.S.	Nov 2011	NSF C-130	Stith and Rogers [22]	Pla, Pla.Agg., Sph, Col, Col.Agg., Den, Gra, Agg
POSIDON	Western Pacific	Oct 2016	NASA's WB-57	Gao et al. [23]	Ros, Bud, Col, Den, Gra, Agg
PREDICT	Atlantic	Aug - Sep 2010	Gulfstream-V	Montgomery et al. [24]	Pla, Pla.Agg., Sph, Den
SEAC4RS	Gulf of Mexico	Sep 2013	SPEC-Learjet	Toon et al. [25]	Pla, Pla.Agg., Col.Agg., Den, Gra, Agg

Note. ATTREX (Airborne Tropical Tropopause Experiment); IDAEAS-4 (Instrument Development and Education in Airborne Science phase 4); POSIDON (Pacific Oxidants, Sulfur, Ice, Dehydration, and Convection Experiment); PREDICT (Pre-Depression Investigation of Cloud-Systems in the Tropics); SEAC4RS (Studies of Emissions and Atmospheric Composition, Clouds \ Climate Coupling by Regional Surveys).

Table 2. Details of the ice crystal particle shape dataset

Category	Number of images	Description
Plate	317	Like a thin, flat plate structure, ice crystals are flat in shape.
Plate Aggregate	269	Aggregates formed by combining multiple plate ice crystals together.
Bullet Rosette	505	Ice crystals show a slender branching structure, similar to the shape of a rose.
Budding Rosette	625	Combination of short columns showing an embryo structure.
Sphere	1053	Mostly spherical or almost spherical particles, including cloud drops and raindrops.
Column	453	Rectangular shape, one axis longer than the other, no signs of aggregate.
Column Aggregate	248	Aggregates formed by combining multiple column ice crystals together.
Dendritic	200	No further ice crystals are condensing on the branch corners, and the six branch formations are spaced uniformly apart.
Graupel	388	As they descend, liquid water droplets collide with ice crystals to generate particles of a specific density.
Aggregate	341	A complex structure created by combining several ice crystals.
Irregular	943	Arbitrary shape with no distinguishing characteristics.
Total	5342	

Natural ice particles have a wide variety of shapes and it is difficult for people to group them in a small number of categories. However, as demonstrated by in-lab tests and aerial observations, specific shapes are identifiable and commonly transpire in clouds. Lindqvist et al. [2] classified ice crystals from CPI into eight types: plate, bullet, column, irregular, bullet rosette, budding rosette, plate aggregate, and column aggregate. The classification proposed by Praz et al. [10] classified ice crystal images into six categories: sphere, column, plate, compact particle, bullet rosette and aggregate. Xiao et al. [17] established the Ice Crystals Database in China (ICDC), which classified ice crystal images into ten types: bullet rosette, budding rosette, aggregate, plate, sector plate, long column, short column, hollow column, sphere, and small irregular. Nine categories were created by Przybylo

et al. [18] based on the ice crystal particles images by CPI: rimed, aggregate, bullet rosette, budding rosette, column, irregular, blur, plate, and sphere. In addition, CPI probe can record particle images with a high resolution of $2.3\mu\text{m}$ and 256 gray levels, compared to the pixel of the optical array probes (2D-S: $10\mu\text{m}$; CIP: $25\mu\text{m}$). However, due to the CPI instrument itself and the external detection environment, there is still uncertainty that the captured ice crystal images may be out of focus and oversized [26]. Eleven categories are identified by careful examination of the ice crystals: dendritic (Den), sphere (Sph), graupel (Gra), aggregate (Agg), bullet rosette (Ros), budding rosette (Bud), column (Col), column aggregate (Col.Agg.), plate (Pla), plate aggregate (Pla.Agg.), and irregular (Irr). A representative image of each shape classification is shown in Fig. 1. Compared with other classifiers, the number of classified categories increases. It can bring the following benefits. First, by increasing the number of classification categories, different types of ice crystals can be distinguished more finely, thus providing more detailed and comprehensive information. Secondly, different shapes of ice crystals reflect different microphysical processes in the atmosphere [27]. Understanding the mechanisms of ice crystal formation, growth, and their interactions with climate factors including clouds, precipitation, and atmospheric circulation can be improved by improving the classification of ice crystal shapes.

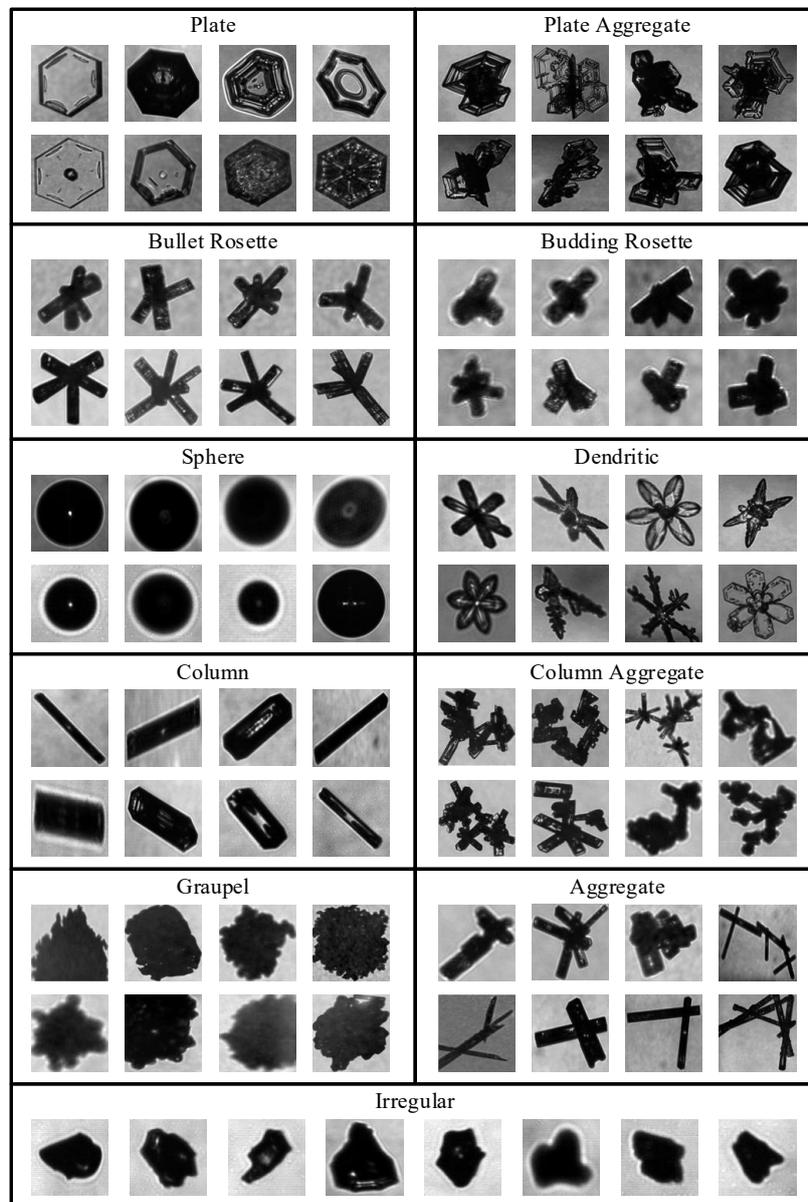


Fig. 1. Typical example diagram of ice crystal particle shape categories in this study

3 Methodology

In this section, we present the CNN classification method used in the experiment for the type of ice crystal particles measured by the CPI probe. Firstly, the implementation of the CNN technique is studied in depth. Then the image preprocessing is described in detail. Finally, the training details are described.

3.1 Network Architecture

CNN is a feed forward neural network category deep learning system that employs the convolutional layer to extract picture features, which are then used in the output layer to perform classification tasks. A typical CNN model architecture typically comprises input layer, hidden layers, and output layer. The hidden layer can be further divided into multiple layers, including the convolutional layer, fully connected layer, activation function, pooling layer, and more [28]. The convolutional layer is the most important component of CNN, and its main procedure is convolution operation. Its operational formula is as follows:

$$f(t) * g(t) = \int_{-\infty}^{+\infty} f(x) * g(t-x) dx \quad (1)$$

In the past decades, CNNs have found successful applications across various research domains [29-32]. These studies demonstrate that CNN algorithms are highly relevant techniques for image analysis. In 2012, Krizhevsky et al. [33] proposed a classical CNN architecture, AlexNet, which won the image classification competition and gained significant attention for CNNs. Subsequently, researchers have proposed numerous new models, including VGGNet [34] and ResNet [35]. While all these prevalent CNN architectures are suitable for classification tasks, each model possesses distinctive attributes. VGGNet, as a classical model, is an important improvement of Alexnet. It reflects the trend of increasing network depth while replacing the use of a 7×7 large convolutional kernel with a 3×3 small convolutional kernel. This improvement reduces the parameter footprint and improves the model's nonlinear processing capability, thus enhancing its discriminative power.

The CNN model used during this work is designed with reference to the VGGNet architecture. When applied to image classification, CNN algorithm consists of two parts: a feature extractor and a classifier. Both have a large number of learnable parameters, which will be continuously optimized to update the network with the required parameters during the learning phase using back propagation algorithm. In this paper, the feature extractor of the CNN model comprises three fundamental layers: the convolutional layer, the activation function layer, and the pooling layer. The convolutional layer serves as the central component of the CNN, employing a series of convolutional kernels to execute convolution operations on the input image, thereby extracting features. The smaller convolution kernels require fewer parameters and computations, while larger convolution kernels capture more complex image information. The activation function layer nonlinearly transforms the output from the convolutional layer, thereby enhancing the model's expressiveness. After normalizing and applying the activation function, each convolution kernel produces a set of feature maps for input image. These feature maps are then downsampled with a 2×2 pooling layer. The pooling layer's primary goal is to process less data overall while keeping relevant information intact and speeding up network training. Pooling is an aggregation operation that reduces the target resolution and aggregates statistics for features at different locations. Two common types of pooling operations are mean pooling and maximum pooling. Mean pooling has a smoothing effect on the extracted features, while maximum pooling highlights features. The feature extraction part of the CNN model in this paper repeats convolution layer, batch normalization operation, activation function layer and pooling layer until each feature map is reduced to a 1×1 size. Among them, the convolutional layers utilize 3×3 convolutional kernels with stride $S=1$, same padding and ReLU activation. The ReLU activation function is expressed by equation (2). The number of convolution kernels increases throughout the convolution part of the model, from 8 to 16, 32, 64, 128, and finally 256. The pooling layers conduct max pooling with a stride of $S=2$. In the classifier segment of the network, two fully connected layers are utilized to reduce the dimensionality from 512 to 11. Then, the Softmax layer normalizes the output of the fully connected layers, generating probability distributions that interpret the model's output as probabilities for each category. ReLU activation function is also employed as an activation function for the fully connected layer. The specific model is depicted in Fig. 2.

$$f(x) = \text{Max}(0, x) . \quad (2)$$

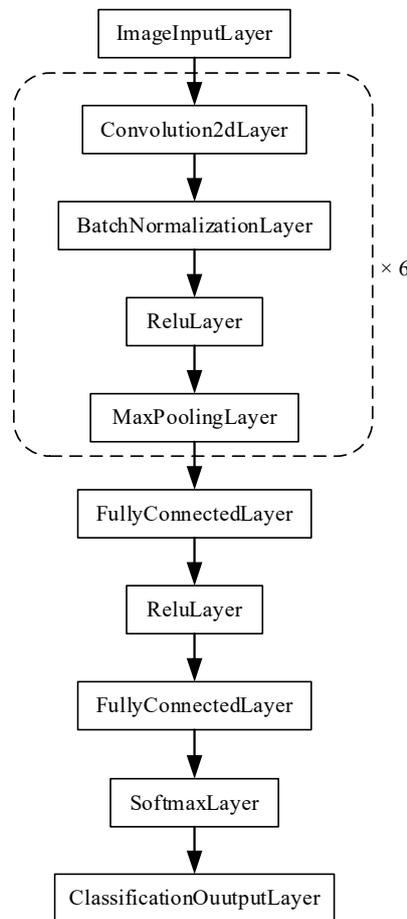


Fig. 2. Network architecture of CPI classification model

3.2 Data Preprocessing

(1) Normalizing the input

Normalizing a picture's pixel values into an appropriate range, like $[0, 1]$ or $[-1, 1]$, is essential before feeding the image into a CNN. The performance of the model may be impacted if images with pixel values between 0 and 255 are directly supplied. This can cause gradient explosion or disappearance. He et al. [36] explained the need for fixed input image sizes in CNN models. If the input vector's dimensionality is not fixed, it causes dynamic changes in the network, preventing effective parameter training. Therefore, it is necessary to fix the input image size when using a fully connected layer. To expedite convergence and gradient descent in the CNN model, we applied the same transformation to each dataset segment: resizing ice crystal particle images to 100×100 and normalizing the pixel matrix.

(2) Data augmentation

Data augmentation is an excellent way to reduce CNN generalization errors and helps prevent the network from overfitting. Data augmentation transformations include rotation, scaling, cropping, flipping, and panning. We used the data augmentation operation after 2200 randomly selected images of ice crystal particles and only on the training set. During CNN training, each image undergoes a transformation operation with a 50% probability before being input into the network. This implies that an image may be subjected to multiple transformations during training, or it may undergo no transformations at all.

3.3 Training Details

This work makes use of the Machine Learning and Deep Learning Toolbox with MATLAB R2021B as its experimental operating platform. The system configuration and hardware utilized are AMD Ryzen 3 3300U with Radeon Vega Mobile Gfx@2.10 GHz and 16GB RAM.

The training details of CNN are as follows. First, 2200 images of ice crystal particles, randomly selected from the dataset, are divided into training and test sets at a ratio of 7:3. This ratio division reduces the possibility of network overfitting to the data and allows for accurate testing of generalization. In order to optimize the network parameters, we employed the Sgdm optimizer, a deep learning gradient descent algorithm. Sgdm introduces first-order momentum to mitigate the issue where traditional stochastic gradient descent (SGD) encounters zero gradients at local optima, thereby enabling continuous updates and addressing the problem of excessive oscillations in oscillation amplitude [37]. We utilized a momentum parameter [38] with a decay of 0.9 to get suitable parameters for the final convolutional layer. The learning rate was set to $\eta=0.01$, which was halved every time 10 epochs were passed. The gradient was computed using a small batch of 16 samples, and the epochs to 25. Every 15 iterations, the model is assessed on the validation set to track its improvement and determine whether overfitting has occurred. Lastly, the model is assessed using the Section 4 assessment parameters.

4 Baseline of Classification Algorithms used for Comparison

Since both support vector machine (SVM) and back propagation (BP) neural network exhibit excellent learning performance and are capable of effectively addressing high-dimensional and nonlinear problems, they are the preferred methods for pattern recognition. Therefore, we conducted a performance comparison between CNN and SVM and BP neural networks in shape recognition models. Both techniques were trained and tested on identical datasets. The specifics of the compared models are delineated below.

4.1 Support Vector Machine

SVM is a general feed-forward network model and is one of the most influential methods in supervised learning. Its main objective is to find a hyperplane that accurately separates different types of datasets. The optimal hyperplane, also referred to as the best hyperplane, maximizes the distance between the nearest data points from each dataset. In cases where the datasets are not linearly separable, the samples from the original space can be projected by SVM into a higher-dimensional feature space, where they can be separated linearly.

For each object in the ice crystal particle images, they require a set of feature vectors (see Table 3). The creation of the SVM model primarily hinges on parameter selection, including the kernel function and penalty factor. In this paper, we compared experiments to construct SVM classification recognition model by applying the libsvm toolbox developed by Zhijian Lin and his lab in Matlab platform at National Taiwan University. The SVMcgForClass function in the libsvm toolbox is applied to automatically search for the optimal penalty factor c and γ parameters for the SVM model. Since determining the kernel function lacks a clear methodology, we assessed the performance of SVM using various kernel functions (linear, polynomial, radial basis function, and sigmoid) in identifying 11 types of ice crystals. Ultimately, we compared the model constructed with a polynomial kernel function to the CNN model.

Table 3. Explanation of every feature utilized in the SVM and BP neural network training and testing

Feature	Component
Aspect ratio	Ratio of the longest longitudinal axis to the longest horizontal axis of ice crystal particles.
Area ratio	Ratio of ice crystal particle area to specific area.
Rectangularity	Ratio of the area of particles to the smallest external rectangle of particles.
Circularity	The degree to which ice crystal particles tend to be round.
Normalized descriptor	An important way to describe the shape of the particle, when the number of dimensions is greater than 20, the normalization factor is already close to 0, so take the first 20 dimensions of the normalized descriptor features.
Local binary pattern	Capable of extracting local texture features of ice crystal particle images, the dimension of feature parameters for each particle is 256.

4.2 BP Neural Network

BP neural network, a type of feed-forward neural network, adjusts network weights by back-propagating the error to calculate the error in each hidden layer. It consists of three or more layers and offers a simple structure and easy training compared to other networks. The training process involves two steps: signal forward propagation and error backward propagation. Signal forward propagation involves feeding particle feature parameters into the network as input, propagating activation values through the input, hidden, and output layers, and comparing the output values with expected values based on the error criterion. The network parameters are adjusted in each layer through repeated training iterations until the output error reaches the desired level.

Matlab neural network toolbox is used by us to construct BP neural network model. To ensure parity in comparison with the CNN model for classification, we applied the same dataset to train the BP neural network as used for the CNN model. The BP neural network was trained on the identical set of features as the SVM (see Table 3). We also compared the performance of the BP neural network models constructed by `trainscg`, `traingdx`, `traingdm` and `traingd`, and selected the best training function to construct the classification model for comparison with the CNN model.

Lastly, the BP neural network's configuration settings are set up as follows. The target expectation error is 0.01, the maximum training times is 1000, the hidden layer's activation function is a `tansig` function, the output layer's activation function is a `logsig` function, the `trainscg` training function is used for training, and the starting learning rate is 0.05.

4.3 Performance Measure

Assessing the generalization performance of automatic image classifiers necessitates not only reliable and feasible experimental estimation methods but also evaluation criteria to gauge the model's generalization capability. Employing various performance measures can often yield divergent assessment outcomes. Hence, we employed multiple performance metrics to assess various facets of the model's performance. The subsequent section provides an overview of the model's performance metrics.

In machine learning, FP denotes judging counterexamples incorrectly as positive examples; TP denotes judging positive examples correctly as positive examples; FN denotes judging positive examples incorrectly as counterexamples; and TN denotes judging counterexamples correctly as counterexamples.

(1) Confusion matrix

The confusion matrix can be used as a basic visualization tool for classification problems and can give a better understanding of the errors in the classification. It is easy to observe where there are mistakes in the confusion matrix as all of the accurate predictions are situated on its diagonal.

(2) Accuracy

Accuracy is the most basic classification model performance evaluation metric. The percentage of correctly identified samples to the total number of samples is indicated.

$$A_c = \frac{TP + TN}{TP + FP + TN + FN} . \quad (3)$$

(3) Precision

Precision is the proportion of ice crystal particles correctly classified as the i -th category as a percentage of the total number of particles identified the i -th category ($i = 1, 2, 3, \dots, 11$).

$$P = \frac{TP}{TP + FP} . \quad (4)$$

(4) Recall

Recall is the proportion of the number of ice crystal particles correctly classified as the i -th category to the total number of particles in i -th category ($i = 1, 2, 3, \dots, 11$).

$$R_c = \frac{TP}{TP + FN} . \quad (5)$$

Precision and recall are complementary metrics. Typically, when precision is high, recall tends to be low, and vice versa. It is usually only possible to have high precision and recall in some simple classification tasks.

(5) F1 score

F1 score is the summed average of the precision and recall rates.

$$F_1 = \frac{2TP}{2TP + FN + FP} . \quad (6)$$

(6) Macro-Precision, macro-Recall and macro-F1 score

The multi-classification problem is split into n pairwise binary classification problems, so that n confusion matrices can be obtained, and Macro-Precision, macro-Recall and macro-F1 score are calculated based on the average values of each category P and R_c . These three performance metrics are used to measure the overall performance of model. The calculation formulas are as follows:

$$marco_P = \frac{1}{n} \sum_{i=1}^n P_i . \quad (7)$$

$$marco_R = \frac{1}{n} \sum_{i=1}^n R_{ci} . \quad (8)$$

$$marco_F_1 = 2 \frac{marco_P * marco_R_c}{marco_P + marco_R_c} . \quad (9)$$

(7) Standard deviation

Standard deviation of the recognition accuracy is used to assess the model's robustness and is calculated as:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (10)$$

where x_i ($i = 1, 2, 3, \dots, 11$) indicates the recognition accuracy of the model for i -th category particle and \bar{x} indicates the average recognition accuracy.

5 Results and Discussion

In this section, we initially present the results of the experimental comparison among three models: CNN, SVM, and BP neural network. Subsequently, we briefly discuss the classification process of the CNN model.

5.1 Performance Comparison of Different Classification Algorithms

Every model that achieved the highest accuracy rating during training was saved and assessed on the test set. The highest accuracy of CNN model, SVM model, and BP model were 95.45%, 83.94%, and 89.55%, respectively. We also examined the prediction outcomes of these three training models on the test set and give the results of the experimental comparison.

Fig. 3 shows the accuracy and cross-entropy loss values obtained from the CNN model network during training and testing. The training dataset comprised 1540 ice crystal particle images, and the training process spanned 2400 iterations. The CNN model achieved an accuracy of 95.45% with a loss function of 0.08, showing little variance between each run. The training duration was 13 minutes and 18 seconds. As illustrated in Fig. 3, the accuracy and loss values stabilized when the epoch number approached 12. Fig. 4(a) displays the confusion matrix of the ice crystal particle recognition results of the CNN model on test set. This matrix exhibits the total number of images for each possible classified particle, revealing that most types of ice crystal shapes are distinguished quite well. It's evident from Fig. 4(a) that the model exhibits minimal prediction errors in the classification of bullet rosette, column, dendritic, graupel, irregular, plate, budding rosette, and sphere. However, the CNN model faces challenges in differentiating between column aggregate and aggregate. Increasing the number of samples could potentially enhance the network's ability to distinguish between these two categories.

Fig. 4(b) and Fig. 4(c) depict the confusion matrices for the recognition results on the test set using the SVM and BP neural network models, respectively. Both models achieved their best classification results for the "sphere" category when trained and tested on the same dataset. As previously mentioned, the CNN model achieved the highest accuracy of 95.45%. Conversely, the SVM model exhibited the lowest accuracy of 83.94%, which is lower than both the CNN and BP neural network models. This discrepancy can be attributed to the strong dependence of SVM on the choice of kernel function and parameter tuning. Selecting an inappropriate kernel function or mis-tuning the parameters can lead to decreased classifier accuracy. In contrast, the BP neural network and CNN models are more intuitive and simpler to tune parameters, as they automatically optimize network weights and biases through the back propagation algorithm. They also have stronger expressive power in handling non-linear problems. We also evaluated macro-Precision, macro-Recall, macro-F1 score, and the standard deviations of recognition accuracy for the three models at the highest accuracy during testing (see Table 4). The CNN model exhibited better performance in terms of macro-Precision, macro-Recall, and macro-F1 score, all exceeding 95%, whereas the SVM and BP models did not exceed 90%. Additionally, the CNN model demonstrated a lower standard deviation of recognition accuracy compared to the SVM and BP models, indicating better model robustness. In summary, our proposed CNN model is effective and has superior performance in the ice crystal particle classification task.

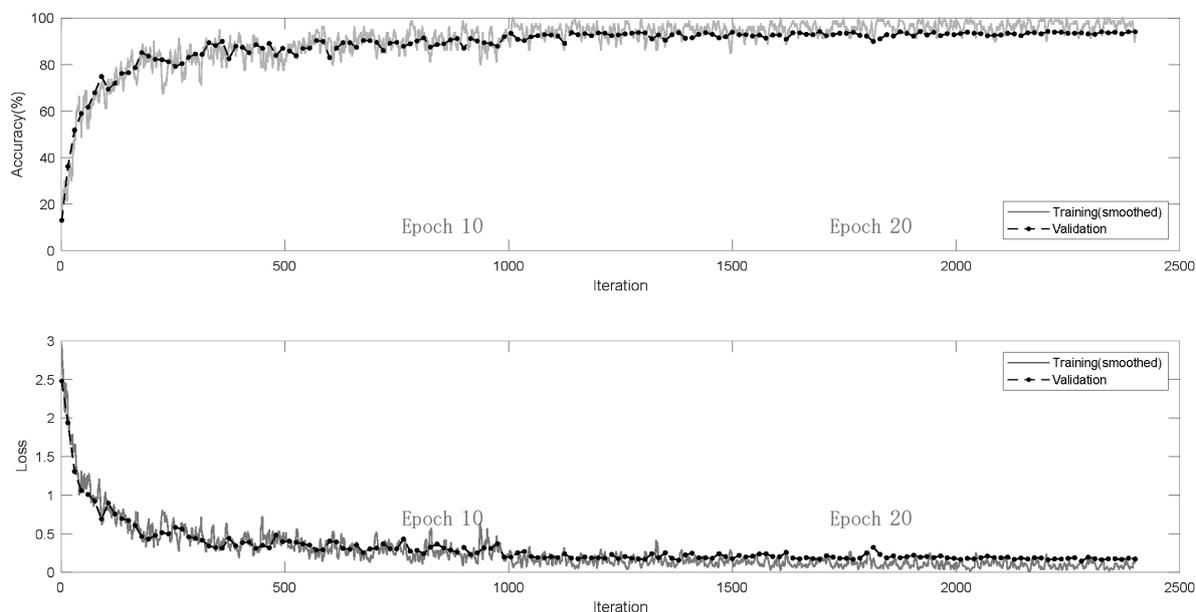
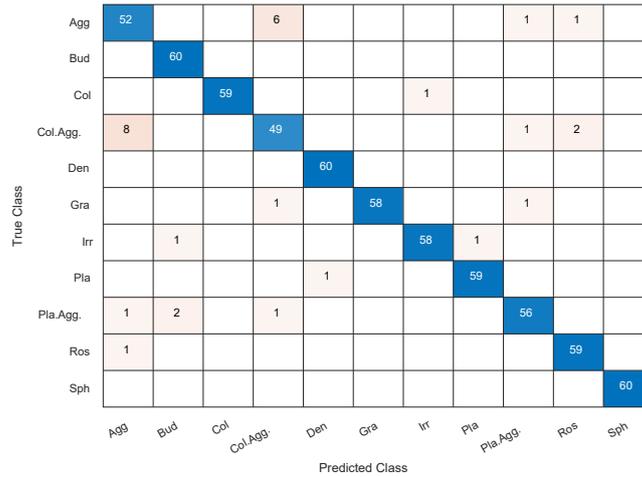
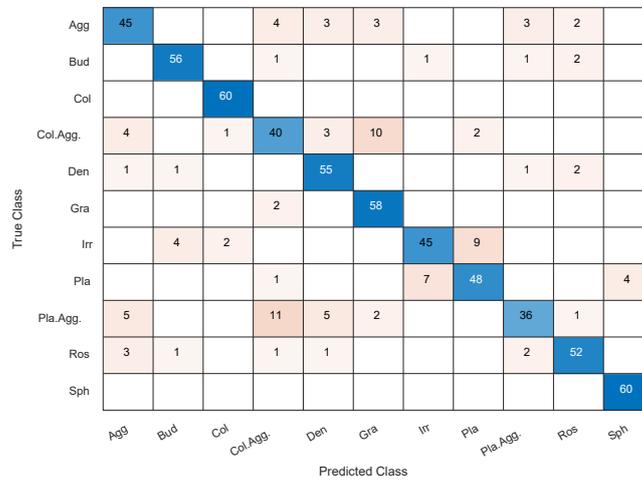


Fig. 3. The model underwent training for 25 epochs, comprising a total of 2400 iterations

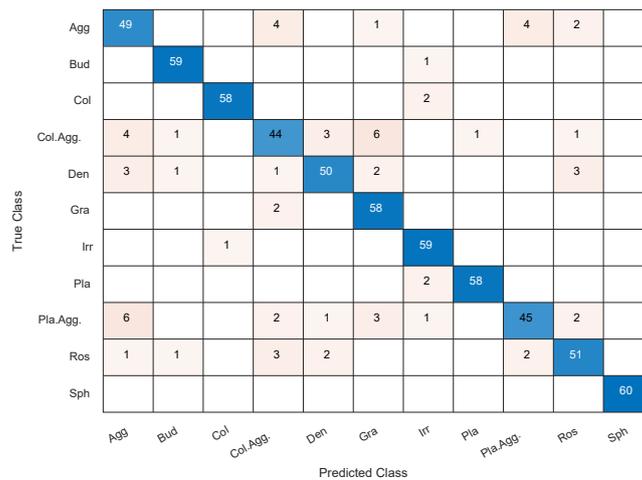
[(top) The average value is verified every 15 iterations (black dots) over 25 epochs with an accuracy of 95.45%. (Bottom) The loss function evaluation shares the same parameters with a final value of 0.08.]



(a) CNN



(b) SVM



(c) BP neural network

Fig. 4. Confusion matrices of ice crystal particle recognition results on the test set

(The true and predicted labels are on the vertical and horizontal axes, respectively, and the correct classification is located on the diagonal line.)

Table 4. Comparison of macro_Precision, macro_Recall and macro_F1 and standard deviation

Method	<i>macro_P</i> (%)	<i>macro_R</i> (%)	<i>macro_F₁</i> (%)	σ
CNN	95.47	95.45	95.46	0.0601
SVM	83.91	83.94	83.93	0.1361
BP	89.61	89.55	89.58	0.1006

We also used the three models to perform additional assessments of precision, recall, and F1 score for every class of ice crystal particles (see to Table 5). The sphere and column shapes performed well across all three models, likely due to their distinct and easily recognizable features. Conversely, column aggregate and aggregate were the most challenging to classify among three models, likely because of the similarities in their complex structures. It is worth noting that the CNN model can reach 100% precision for classifying sphere, column and graupel. Except for the irregular's recall, the recall and F1score of CNN model are the highest values of the three models, and the performance measures for most categories are higher than 95%. Moreover, it is evident that the CNN model achieves the maximum F1 score for every category. To assess the classification model more effectively, the F1 score can take precision and recall into account. The aforementioned findings demonstrate that the suggested CNN model outperforms conventional classification techniques that call for manual feature extraction in terms of generalization and learning with small data.

Table 5. Experimental evaluation of the three models

Category	Method	<i>P</i> (%)	<i>R_c</i> (%)	<i>F₁</i> (%)
Plate	CNN	98.33	98.33	98.33
	SVM	81.36	80.00	80.67
	BP	98.31	97.00	97.48
Plate Aggregate	CNN	94.92	93.33	94.12
	SVM	83.72	60	69.90
	BP	88.24	75.00	81.08
Bullet Rosette	CNN	95.16	98.33	96.72
	SVM	88.14	86.67	87.39
	BP	86.44	85.00	85.71
Budding Rosette	CNN	95.24	100	97.56
	SVM	90.16	91.67	90.91
	BP	95.16	98.33	96.72
Sphere	CNN	100	100	100
	SVM	93.75	100	96.77
	BP	100	100	100
Column	CNN	100	98.33	99.16
	SVM	95.24	100	97.56
	BP	98.31	96.67	97.48
Column Aggregate	CNN	85.96	81.67	83.76
	SVM	66.67	66.67	66.67
	BP	78.57	73.33	75.86
Dendritic	CNN	98.36	100	99.17
	SVM	82.09	91.67	86.61
	BP	89.29	83.33	86.21
Graupel	CNN	100	96.67	98.31
	SVM	79.45	96.67	87.22
	BP	82.86	96.67	89.23
Aggregate	CNN	83.87	86.67	85.25
	SVM	77.59	75.00	76.27
	BP	77.78	81.67	79.67
Irregular	CNN	98.31	96.67	97.48
	SVM	84.91	75.00	79.65
	BP	90.77	98.33	94.40

5.2 Discussion

To better understand the classification process of the CNN model, we compared the activation region with the original image to check the activation and discover the features learned by the network, and the activation of the CNN are shown in Fig. 5. By visualizing the activation of CNN, we can observe the sensitivity of the network to different features at different levels. Fig. 5. demonstrates that the shallow channels (Conv_1-4 , Conv_2-8 and Conv_3-16) primarily learn simple features such as edges. These simple features can effectively represent different particle shapes. In contrast, the deep channels (Conv_4-32, Conv_5-64 and Conv_6-128) learn more abstract and complex features, exhibiting only local characteristics. Deeper layers construct their features by combining features from shallower layers. Consequently, the CNN's convolutional layer effectively extracts local features from images. Through the integration of pooling and fully connected layers, the CNN progressively learns increasingly abstract feature representations, forming a layered feature extraction mechanism. This mechanism enables CNNs to gradually acquire features at different levels within an image, making them valuable for image classification tasks.

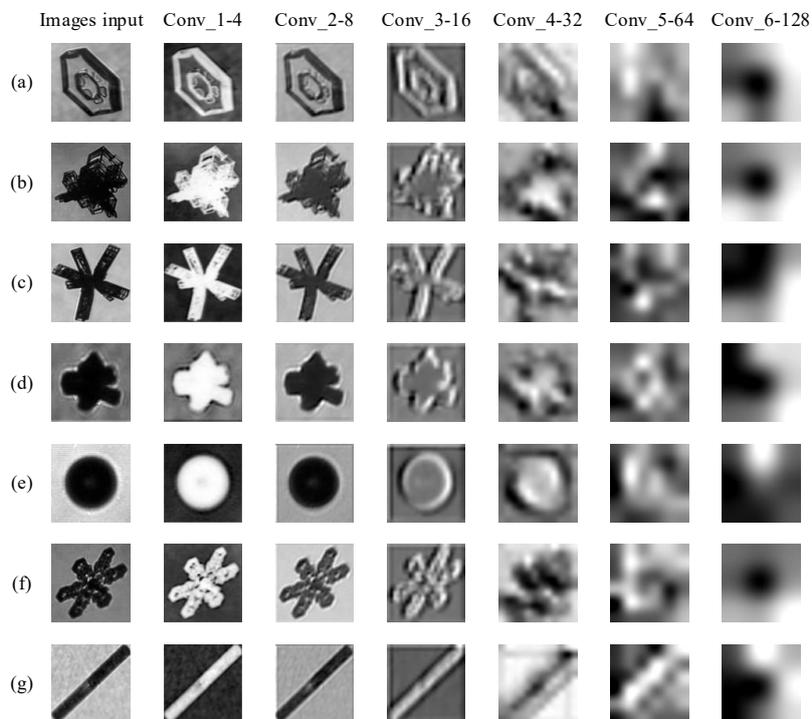


Fig. 5. Visualization of feature maps in CNN model, denoted Conv_1-4 (channel 4 on Conv_1), Conv_2-8, Conv_3-16, Conv_4-32, Conv_5-64, and Conv_6-128

In addition, when server resources are not available, the issue of extended training time for CNN may become particularly pronounced. Compared with the TL-ResNet152 model and TL-EfficientNet-b6 model proposed by Xiao et al. [17] and Wu et al. [19], the CNN model constructed in this paper has a simpler structure, requiring less time for training. It can classify more different types of ice crystals, and achieve a good level of classification accuracy.

6 Conclusion and Outlook

An automatic ice crystal particle shape classifier based on CNN was created to derive particle habit from CPI that is routinely used aboard research airplane in cloud observation community. The proposed CNN model clas-

sifier was a 30-layer network and built with small 3×3 convolutional kernels. For developing and testing the classifier, data collected over continental regions of China and North America, and the Pacific and Atlantic Oceans by CPI were used to establish an ice crystal habit dataset, which consists of eleven shape categories containing 5,342 images. Using this dataset, we compared the performance of the CNN classifier with that of traditional machine learning models such as SVM and BP. The CNN achieved an accuracy of 95.45%, while SVM and BP neural network achieved accuracies of 83.94% and 89.55%, respectively. The CNN model performed the best and obtained the highest F1 score for each category. These experimental results demonstrate the CNN model's effectiveness and superiority in ice crystal shape classification tasks. Compared to the traditional methods, the CNN model can better learn essential features in ice crystal particle images, enabling improved discrimination of different ice crystal types.

Furthermore, we conducted an in-depth analysis of the CNN model to gain insights into its functioning in ice crystal shape classification. Our findings indicate that the CNN model can effectively capture features such as shape, texture, and structure of ice crystals, which are crucial for accurate classification and identification. Through visualization and feature mapping analysis, we further confirmed the CNN model's efficacy and interpretability in ice crystal shape classification. Overall, the proposed automatic particle image classification method achieves better classification results.

However, our algorithm still requires improvements in recognizing the shapes of complex ice crystal particle aggregates. Future research should focus on refining and optimizing the CNN model proposed in this study to enhance its accuracy in recognizing complex aggregates' shapes. This will enable better discrimination of various ice crystal shapes and facilitate broader applications in meteorological and climate research, fostering deeper understanding and analysis of ice crystal particles.

7 Acknowledgement

This research was funded by the Instrument and Equipment Function Development Project of Chinese Academy of Sciences (E3661716), the Innovation Foundation of CPML/CMA (2023CPML-B06), the Open Project Fund of Beijing, Nanjing Joint Institute for Atmospheric Sciences (BJG202302), the Technical Support Talent of Chinese Academy of Sciences in 2022 (E3661416), the National Natural Science Foundation of China (41775166 and 41705142), and the Beijing Information Science & Technology University Project (S2326188).

References

- [1] M. Wendisch, P. Yang, P. Pilewskie, Effects of ice crystal habit on thermal infrared radiative properties and forcing of cirrus, *Journal of Geophysical Research: Atmospheres* 112(D8)(2007).
- [2] H. Lindqvist, K. Muinonen, T. Nousiainen, J. Um, G.M. McFarquhar, P. Haapanala, R. Makkonen, H. Hakkarainen, Ice-cloud particle habit classification using principal components, *Journal of Geophysical Research: Atmospheres* 117(D16) (2012) 1-12.
- [3] A.J. Baran, From the single-scattering properties of ice crystals to climate prediction: A way forward, *Atmospheric Research* 112(2012) 45-69.
- [4] W. Wu, G.M. McFarquhar, On the impacts of different definitions of maximum dimension for nonspherical particles recorded by 2D imaging probes, *Journal of Atmospheric and Oceanic Technology* 33(5)(2016) 1057-1072.
- [5] J. Um, G.M. McFarquhar, Dependence of the single-scattering properties of small ice crystals on idealized shape models, *Atmospheric Chemistry and Physics* 11(7)(2011) 3159-3171.
- [6] H. Letu, H. Ishimoto, J. Riedi, T.Y. Nakajima, L. C.-Labonnote, A.J. Baran, T.M. Nagao, M. Sekiguchi, Investigation of ice particle habits to be used for ice cloud remote sensing for the GCOM-C satellite mission, *Atmospheric Chemistry and Physics* 16(18)(2016) 12287-12303.
- [7] H. Letu, T.M. Nagao, T.Y. Nakajima, J. Riedi, H. Ishimoto, A.J. Baran, H. Shang, M. Sekiguchi, M. Kikuchi, Ice cloud properties from Himawari-8/AHI next-generation geostationary satellite: Capability of the AHI to monitor the DC cloud generation process, *IEEE Transactions on Geoscience and Remote Sensing* 57(6)(2018) 3229-3239.
- [8] A. Korolev, B. Sussman, A technique for habit classification of cloud particles, *Journal of Atmospheric and Oceanic Technology* 17(8)(2000) 1048-1057.
- [9] R.P. Lawson, B.A. Baker, P. Zmarzly, D. O'Connor, Q. Mo, J.-F. Gayet, V. Shcherbakov, Microphysical and optical properties of atmospheric ice crystals at South Pole Station, *Journal of Applied Meteorology and Climatology* 45(11) (2006) 1505-1524.

- [10] C. Praz, S. Ding, G.M. McFarquhar, A. Berne, A versatile method for ice particle habit classification using airborne imaging probe data, *Journal of Geophysical Research: Atmospheres* 123(23)(2018) 13472-13495.
- [11] M. Huang, H. Lei, An improved Holroyd cloud particle habit identification method and its application, *Acta Meteorologica Sinica* 78(2)(2020) 289-300.
- [12] H.N. Dong, R.L. Jiao, M.S. Huang, Method of cloud particle shape automatic recognition based on BP neural network optimized by simulated annealing algorithm, *Modern Electronics Technique* 45(12)(2022) 143-148.
- [13] G. Touloupas, A. Lauber, J. Henneberger, A. Beck, A. Lucchi, A convolutional neural network for classifying cloud particles recorded by imaging probes, *Atmospheric Measurement Techniques* 13(5)(2020) 2219-2239.
- [14] M. Liao, J. Duan, R. Zhang, X. Zhou, X. Wu, X. Wang, J. Hu, A hypergraph-embedded convolutional neural network for ice crystal particle habit classification, *Intelligent Automation & Soft Computing* 29(2021) 787-801.
- [15] R. Jiao, M. Huang, X. Chang, Phase Recognition of Supercooled Droplet Based on Convolutional Neural Network, in: *Proc. 2021 IEEE 4th International Conference on Computer and Communication Engineering Technology (CCET)*, 2021.
- [16] J. Huo, M. Huang, R. Jiao, Design and Implementation of a Cloud Particle Shape Recognition Algorithm Based on MATLAB, in: *Proc. 2022 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2022.
- [17] H. Xiao, F. Zhang, Q. He, P. Liu, F. Yan, L. Miao, Z. Yang, Classification of ice crystal habits observed from airborne Cloud Particle Imager by deep transfer learning, *Earth and Space Science* 6(10)(2019) 1877-1886.
- [18] V.M. Przybylo, K.J. Sulia, C.G. Schmitt, Z.J. Lebo, Classification of cloud particle imagery from aircraft platforms using convolutional neural networks, *Journal of Atmospheric and Oceanic Technology* 39(4)(2022) 405-424.
- [19] Z. Wu, S. Liu, D. Zhao, L. Yang, Z. Xu, Z. Yang, W. Zhou, H. He, M. Huang, D. Liu, R. Li, D. Ding, Neural network classification of ice-crystal images observed by an airborne cloud imaging probe, *Atmosphere-Ocean* 58(5)(2020) 303-315.
- [20] S. Glienke, F. Mei, *Cloud Particle Imager (CPI) and 3-View Cloud Particle Imager (3V-CPI) Instrument Handbook*, United States, 2020.
- [21] E.J. Jensen, L. Pfister; D.E. Jordan, T.V. Bui, R. Ueyama, H.B. Singh, T.D. Thornberry, A.W. Rollins, G. R.-Shan, D.W. Fahey, K.H. Rosenlof, J.W. Elkins, G.S. Diskin, J.P. Digangi, R.P. Lawson, S. Woods, E.L. Atlas, R.M.A. Navarro, S.C. Wofsy, J. Pottman, The NASA Airborne Tropical Tropopause Experiment: High-altitude aircraft measurements in the tropical western Pacific, *Bulletin of the American Meteorological Society* 98(1)(2017) 129-143.
- [22] J. Stith, D.C. Rogers, Instrument development and education in airborne science, in: *Proc. 13th Symposium on Education*, 2004.
- [23] R.S. Gao, E.J. Jensen, K.H. Rosenlof, E.L. Atlas, M.A. Avery, T.V. Bui, P.R. Colarco, G.S. Diskin, J.W. Elkins, E.J. Hints, D.F. Hurst, P. Lawson, S. Liu, F.L. Moore, L. Pfister, A.W. Rollins, S. Schauffler, M.R. Schoeberl, T.D. Thornberry, R. Ueyama, J. Ulanowski, S. Woods, Overview of the NASA Pacific Oxidants, Sulfur, Ice, Dehydration, and cONvection (POSIDON) Experiment, in: *Proc. AGU Fall Meeting Abstracts*, 2016.
- [24] M.T. Montgomery, C. Davis, T. Dunkerton, Z. Wang, C. Velden, R. Torn, S.J. Majumdar, F. Zhang, R.K. Smith, L. Bosart, M.M. Bell, J.S. Haase, A. Heymsfield, J. Jensen, T. Campos, M.A. Boothe, The Pre-Depression Investigation of Cloud-Systems in the Tropics (PREDICT) experiment: Scientific basis, new analysis tools, and some first results, *Bulletin of the American Meteorological Society* 93(2)(2012) 153-172.
- [25] O.B. Toon, H. Maring, J. Dibb, R. Ferrare, D.J. Jacob, E.J. Jensen, Z.J. Luo, G.G. Mace, L.L. Pan, L. Pfister, K.H. Rosenlof, J. Redemann, J.S. Reid, H.B. Singh, A.M. Thompson, R. Yokelson, P. Minnis, G. Chen, K.W. Jucks, A. Pszenny, Planning, implementation, and scientific goals of the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC4RS) field mission, *Journal of Geophysical Research: Atmospheres* 121(9)(2016) 4967-5009.
- [26] S. Jang, J. Kim, G.M. McFarquhar, S. Park, S. Han, S.S. Lee, C.H. Jung, H. Jung, K.-H. Chang, W. Jung, J. Um, The Impacts of Single-Scattering and Microphysical Properties of Ice Particles Smaller Than 100 μm on the Bulk Radiative Properties of Tropical Cirrus, *Remote Sensing* 14(13)(2022) 3002.
- [27] A.V. Korolev, M.P. Bailey, J. Hallett, G.A. Isaac, Laboratory and in situ observation of deposition growth of frozen drops, *Journal of Applied Meteorology and Climatology* 43(4)(2004) 612-622.
- [28] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen, Recent advances in convolutional neural networks, *Pattern recognition* 77(2018) 354-377.
- [29] P.S. Thakur, T. Sheorey, A. Ojha, VGG-ICNN: A Lightweight CNN model for crop disease identification, *Multimedia Tools and Applications* 82(1)(2023) 497-520.
- [30] Y. Hamid, S. Elyassami, Y. Gulzar, V.R. Balasaraswathi, T. Habuza, S. Wani, An improvised CNN model for fake image detection, *International Journal of Information Technology* 15(1)(2023) 5-15.
- [31] M. Altun, H. Gürüler, O. Özkaraca, F. Khan, J. Khan, Y. Lee, Monkeypox detection using CNN with transfer learning, *Sensors* 23(4)(2023) 1783.
- [32] A.L.C. Ottoni, R.M. de Amorim, M.S. Novo, D.B. Costa, Tuning of data augmentation hyperparameters in deep learning to building construction image classification with small datasets, *International Journal of Machine Learning and Cybernetics* 14(1)(2023) 171-186.
- [33] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: *Proc. 25th International Conference on Neural Information Processing Systems*, 2012.

- [34] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. <<https://arxiv.org/abs/1409.1556v6>>, 2015 (accessed 04.09.2014).
- [35] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [36] K. He, X. Zhang, S. Ren, J. Sun, Spatial pyramid pooling in deep convolutional networks for visual recognition, IEEE transactions on pattern analysis and machine intelligence 37(9)(2015) 1904-1916.
- [37] S. Ruder, An overview of gradient descent optimization algorithms. <<https://arxiv.org/abs/1609.04747v2>>, 2016 (accessed 15.09.2016).
- [38] I. Sutskever, J. Martens, G. Dahl, G. Hinton, On the importance of initialization and momentum in deep learning, in: Proc. International conference on machine learning, 2013.