Classification of Ice Crystal Habits by Deep Transfer Learning

Yefeng Xu
Beijing Information Science and Technology University

Ruili Jiao (jiaoruili@bistu.edu.cn)
Beijing Information Science and Technology University

Minsong Huang
Chinese Academy of Sciences

Research Article

Keywords: Ice crystal habits, Transfer learning, Deep learning, Line Scan Imager, Area Scan Imager

Posted Date: December 5th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3629696/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.
Classification of Ice Crystal Habits by Deep Transfer Learning

Yefeng Xu¹, Ruili Jiao¹, Minsong Huang²

¹ School of Information and Communication Engineering, Beijing Information Science and Technology University, Beijing 100101, China.
² Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China.

*Corresponding author(s). E-mail(s): jiaoruili@bistu.edu.cn; mission@mail.iap.ac.cn; Contributing authors: yefeng.xu@bistu.edu.cn

Abstract. Ice crystal particle shape is an important factor affecting cloud microphysical processes. Accurately identifying the shapes of ice crystal particle within clouds is a fundamental requirement for calculating various cloud microphysical parameters. In this study, we set up an ice crystal image dataset, encompassing nine distinct habit categories with 8100 images. These images were captured using three probes with varying resolutions: the Cloud Particle Imager (CPI), the Two-dimensional Stereo Probe (2D-S), and the High Volume Precipitation Spectrometer (HVPS). In addition, we introduce a deep convolutional neural network (CNN) based on transfer learning for ice crystal particle shape classification model, TL-AlexNet, which demonstrates the capability to simultaneously classify ice crystal particle habits observed by both the Line Scan Imager and the Area Scan Imager. The results indicate that the TL-AlexNet model could achieve superior performance in ice crystal shapes classification for two types of imagers, and the classification with the accuracy of 97.16%. It is much higher than the traditional shape recognition methods, and has certain application value for Climate and cloud microphysics research.

Keywords: Ice crystal habits, Transfer learning, Deep learning, Line Scan Imager, Area Scan Imager
1 Introduction

Covering more than 50 percent of the Earth's surface, clouds have a crucial influence on the process of balanced radiative energy transfer in the Earth-atmosphere system and are the focus of atmospheric science research (Ohring and Adler, 1978; Ramanathan et al., 1988). As a crucial facet of cloud feature, the microphysical play an important role in cloud parameterisation schemes in weather and climate models (Shupe et al. 2008; Wu and McFarquhar 2016) and in the determination of cloud radiative properties in atmospheric remote sensing (Lohmeier et al. 1997; Young et al. 2000). And through the shape and diameter of the ice crystal particles, various cloud microphysical characteristics such as effective particle radius, cloud phase states, extinction coefficients, liquid and ice water content, and single scattering properties can be calculated (Wu and McFarquhar 2016). Therefore, accurate identification of ice crystal particle shapes within clouds holds significant importance in comprehending the changes in cloud microphysics. Additionally, the shape of ice crystal particles influences their scattering properties, growth rate, and terminal velocity. The scattering properties of ice crystals wield substantial influence over the global climate and radiation balance, and bear great significance in inverse corrections applied through remote sensing techniques such as satellites and radar (Kaufman et al. 1997, Korolev et al. 1999, Baran 2009). Therefore, the study of ice crystal particle shapes in clouds holds immense significance for cloud precipitation physics research and the development of climate models (Baker and Lawson 2006).

In recent years, the rapid progress and widespread utilization of artificial intelligence, particularly deep learning algorithms, have made substantial contributions to advancements in the field of image processing and image classification. Deep learning has demonstrated its superiority over traditional machine learning methods in multiple domains (Hedjazi et al. 2017; Chalapathy and Chawla 2019; Chaganti et al. 2020). Presently, deep convolutional neural network (CNN) has been employed for the task of classifying ice crystal particle images. Xiao et al. (2019) proposed TL-ResNet152, an automated ice crystal habit classification model based on deep CNN using transfer learning, achieving an accuracy exceeding 96% in classifying ice crystals into ten shape categories. Touloupas et al. (2020) applied CNN to categorize cloud particles obtained by a holographic imager into droplets, ice crystal particles, and artifacts with 96.8% accuracy. Liao et al. (2021) proposed a CNN embedded with hypergraphic convolutional module, Hy-INet, which achieved a high accuracy of 98.08% in a ten-category classification task for ice crystals. Jaffeux et al. (2022) devised two models to classify ice crystal images acquired by the Precipitation Imaging Probe (PIP) and the Two-dimensional Stereo (2D-S) Probe into 6 and 9 classes using CNN models, respectively, and obtained 85.44% and 91.02% accuracy. Zhang et al. (2023) introduced a CNN-based classification method for 2D-S cloud particle images, achieving an average accuracy of 97% in eight cloud particle habits classification. All of these studies have shown that CNN not only enhances generalization capability but also requires minimal image preprocessing, enabling direct utilization of ice crystal images as input to automatically extract features from the data and make more reliable predictions.

However, in the realm of current ice crystal habit research, the images utilized for classifying ice crystal shapes are predominantly captured by a particular probe, and there are few automatic ice crystal shape classification models that can be applied to
both Line Scan Imager (e.g., 2D-S, PIP, and HVPS, etc.) and Area Scan Imager (e.g., CPI). Chen et al. (2023) proposed training and evaluating CNN model using ice crystal images from Line Scan Imager and Area Scan Imager as a dataset with an accuracy of over 90%. To the best of our knowledge, this model is the first and only CNN model that can simultaneously classify ice crystal images from different types of instruments and achieve good classification results. However, the model has limited classification accuracy and only ice crystal images sampled by CPI and 2D-S probes were used as datasets for model training. In this paper, we manually constructed an ice crystal database containing 9 distinct habit categories with a total of 8100 images. The images are observed by 2D-S, HVPS and CPI probe over continental regions of China and North America, and the Pacific and Atlantic Oceans. Then, building upon the CNN network structure and employing a transfer learning method, this paper proposes an automatic ice crystal classification model named TL-AlexNet. The model enables automatic classification of ice crystal particles measured by both types of instruments, Line Scan Imager and Area Scan Imager, avoiding the need for tedious manual extraction of features (e.g., area, maximum size, roundness, and geometrical feature ratios of ice crystals, etc.) from the images. The subsequent sections of this paper are organized as follows: Section 2 provides a brief overview of the three detection devices and the datasets used. Section 3 introduces the basic concepts of CNN and transfer learning, along with a description of the experimental platforms, training processes, and the method for evaluating model performance. Section 4 conducts a comparative analysis of the classification performance of several representative deep CNN model, selecting the most effective CNN model (TL-AlexNet) for further evaluation, and finally analyses the classification results of the three probe test sets separately. Section 5 summarises and outlooks the research work in this paper.

2 Instrument and dataset

Presently, airborne measurement is considered to be the most direct and effective method of obtaining cloud microphysical characteristics. Baumgardner et al. (2011) and Wendisch and Brenguier (2013) provide an overview of instruments commonly used for cloud microphysical measurement. Airborne cloud particle probes can be categorized into two types based on their image recording methods: the first uses a linear photodiode array, while the second type employs square photodetector arrays to achieve instant imaging of ice crystal particles. The 2D-S (Lawson et al. 2006; Lawson 2011) and HVPS (Lawson et al. 1993) probes utilize linear photodiode array that scan at a rate proportional to particle velocity. Images of ice crystals obtained by both probes are binary due to the utilization of a single threshold (typically 50%) for identifying valid particle images. Comparison to HVPS, 2D-S probe has two sets of vertically orthogonal linear arrays, each equipped with 128 photodiodes. And CPI (Lawson et al. 2001; Glienke and Mei 2020) relies on charge coupled device (CCD) camera. When a particle passes through the sample volume, it triggers a pulsed laser to emit a burst of light, and the resulting particle image is projected onto the digital CCD. This method necessitates pulsed illumination and a triggering system to detect the presence of ice crystal within the sample volume, resulting in discontinuous particle measurement. Although measurement discontinuity is not critical for ice
crystal classification study, it impedes the precise estimation of particle concentration (Baum et al. 2005). In contrast to the 2D-S and HVPS probes, the CPI provides 256 levels of gray scale images, thus offering additional insights into the structure and transparency of particle surface. Table 1 summarizes the cloud particle detection equipment employed in this study and their principal properties.

Table 1. Overview of the cloud particle detection equipment used in this paper.

<table>
<thead>
<tr>
<th>Specifications</th>
<th>CPI</th>
<th>2D-S</th>
<th>HVPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (μm/pixel)</td>
<td>2.3</td>
<td>10</td>
<td>150</td>
</tr>
<tr>
<td>Measurement technique</td>
<td>Square photodiode array</td>
<td>Linear photodiode array</td>
<td>Linear photodiode array</td>
</tr>
<tr>
<td>Measurement range (μm)</td>
<td>2.3-2300</td>
<td>10-1280</td>
<td>150-19200</td>
</tr>
<tr>
<td>Frequency</td>
<td>Depends on pulse illumination</td>
<td>Depends on aircraft speed</td>
<td>Depends on aircraft speed</td>
</tr>
<tr>
<td>Image type</td>
<td>256 levels gray scale</td>
<td>Black and white</td>
<td>Black and white</td>
</tr>
</tbody>
</table>

Table 2. Details of data acquisition for field observation projects. Projects include: ATTREX (Airborne Tropical Tropopause Experiment); IDAEAS-4 (Instrument Development and Education in Airborne Sciencologyphase 4); POSIDON (Pacific Oxidants, Sulfur, Ice, Dehydration, and cONvection Experiment); PREDICT (Pre-Depression Investigation of Cloudsystems in the Tropics); SEAC4RS (Studies of Emissions and Atmospheric Composition, Clouds \ Climate Coupling by Regional Surveys); OLYMPEX (Olympic Mountain Experiment).

<table>
<thead>
<tr>
<th>Campaign</th>
<th>Date</th>
<th>Aircraft</th>
<th>Location</th>
<th>Reference</th>
<th>Major ice crystals</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTREX</td>
<td>Jan - Mar</td>
<td>Global Hawk</td>
<td>Western Pacific</td>
<td>Jensen et al. (2017)</td>
<td>CPI: Pla, Agg., Ros, Bud, Col, Agg</td>
</tr>
<tr>
<td></td>
<td>2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDAEAS-4</td>
<td>Nov 2011</td>
<td>NSFC-C130</td>
<td>North Central U.S.</td>
<td>Jackson et al. (2014)</td>
<td>2D-S: All</td>
</tr>
<tr>
<td>SEAC4RS</td>
<td>Sep 2013</td>
<td>SPEC-Learjet</td>
<td>Gulf of Mexico</td>
<td>Toon et al. (2016)</td>
<td>Sph, Den, Agg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLYMPEX</td>
<td>Nov 2015 -</td>
<td>UND Citation</td>
<td>Olympic Mountains</td>
<td>Houze et al. (2017)</td>
<td>HVPS: All</td>
</tr>
<tr>
<td></td>
<td>May 2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For conducting a study on ice crystal particle shape recognition, a dataset comprising an adequate number of accurately labeled ice crystals is essential. Our dataset of ice crystal images comes from several field observation projects. Ice crystal particle data for CPI were collected from East China and North China, and six field observation projects: ATTREX, IDAEAS-4, POSIDON, PREDICT, SEAC4RS and OLYMPEX. Ice crystal particle information for 2D-S were obtained from Northern China and IDAEAS-4 field project. And the HVPS ice crystal image data were obtained from Eastern China and the OLYMPEX field campaign. Specific information on foreign field observation projects is shown in Table 2. In this paper, an
ice crystal particle shape dataset is constructed manually, and the details of the ice crystal particle shape dataset are shown in Table 3. The dataset contains 8100 representative images of ice crystal particles, with 300 images for each of the different ice crystal habit classes for each probe.

Table 3. Font sizes of headings. Table captions should always be positioned above the tables.

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plate</td>
<td>Ice crystals have a flat shape, similar to a thin and flat plate structure.</td>
</tr>
<tr>
<td>Rosette</td>
<td>Rosette structure with short or long branches.</td>
</tr>
<tr>
<td>Sphere</td>
<td>Solid particle of water that is predominantly round or nearly round, including cloud droplets and raindrops</td>
</tr>
<tr>
<td>Column</td>
<td>Rectangular shape, one axis longer than the other.</td>
</tr>
<tr>
<td>Column Aggregate</td>
<td>An aggregate formed by combining multiple columnar ice crystals together.</td>
</tr>
<tr>
<td>Dendritic</td>
<td>The six branch structures are uniformly distributed, without other ice crystals condensing on the branch corners.</td>
</tr>
<tr>
<td>Graupel</td>
<td>Ice crystals encounter liquid water droplets during their descent and form particles with a certain density.</td>
</tr>
<tr>
<td>Aggregate</td>
<td>A complex structure formed by combining multiple ice crystals together.</td>
</tr>
<tr>
<td>Irregular</td>
<td>Arbitrary shape, no distinctive features.</td>
</tr>
</tbody>
</table>

The size and shape of natural ice crystal particles exhibit intricate diversity, and the process of ice crystal formation and growth is equally complex. Kikuchi et al. (2013) expanded upon the Magono and Lee (1966) classification method for snow crystals, ice crystals, and solid precipitation particles, increasing the number of ice crystal particle shape categories from 80 to 121. However, different researchers have used different numbers and types of shape categories in automatic classifiers for ice crystal images from different probes. For instance, the classification method proposed by Praz et al. (2018) classifies CPI, 2D-S and HVPS probe images into sphere, column, plate, rosette, aggregate and compact particle and the out-of-focus images were added to the classification of 2D-S probes, while the plate crystal was removed from the HVPS classification. Wu et al. (2020) employed nine distinct categories for classifying ice crystal images captured by the Cloud Imaging Probe (CIP), including tiny, sphere, column, needle, irregular, dendrite, graupel, plate, and aggregate. Jaffeux et al. (2022) classified ice crystals from 2D-S probe into nine categories: graupel, fragile aggregate, column, rosette, plate, sphere, capped column, artifact, and aggregate, and particle habits for the Precipitation Imaging Probe (PIP) were classified into six categories: graupel, fragile aggregate, column, rosette, plate, sphere, capped column, and column aggregate. Chen et al. (2022) classified images of ice crystal sampled by 2D-S and CPI probes into ten categories: plate, plate aggregate, irregular, plate, rosette, and rimed aggregate. Considering the differences in resolution between HVPS, 2D-S, and CPI probes, as well as in comparison to other resolution probes (CIP: 25µm; PIP: 100µm), the ice crystals are divided into nine classes: plate (Pla), rosette (Ros), sphere (Sph), column (Col), column aggregate (Col. Agg.), dendritic (Den), graupel (Gra), aggregate (Agg), and irregular (Irr). Figure 1 displays images of each habit captured by the HVPS, 2D-S, and CPI probes.
Fig. 1. Typical examples of ice crystal shapes for the HVPS, 2D-S and CPI probes.

<table>
<thead>
<tr>
<th></th>
<th>HVPS</th>
<th>2D-S</th>
<th>CPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plate</td>
<td><img src="image1.png" alt="Plate Images" /></td>
<td><img src="image2.png" alt="Plate Images" /></td>
<td><img src="image3.png" alt="Plate Images" /></td>
</tr>
<tr>
<td>Rosette</td>
<td><img src="image4.png" alt="Rosette Images" /></td>
<td><img src="image5.png" alt="Rosette Images" /></td>
<td><img src="image6.png" alt="Rosette Images" /></td>
</tr>
<tr>
<td>Dendritic</td>
<td><img src="image7.png" alt="Dendritic Images" /></td>
<td><img src="image8.png" alt="Dendritic Images" /></td>
<td><img src="image9.png" alt="Dendritic Images" /></td>
</tr>
<tr>
<td>Sphere</td>
<td><img src="image10.png" alt="Sphere Images" /></td>
<td><img src="image11.png" alt="Sphere Images" /></td>
<td><img src="image12.png" alt="Sphere Images" /></td>
</tr>
<tr>
<td>Graupel</td>
<td><img src="image13.png" alt="Graupel Images" /></td>
<td><img src="image14.png" alt="Graupel Images" /></td>
<td><img src="image15.png" alt="Graupel Images" /></td>
</tr>
<tr>
<td>Column</td>
<td><img src="image16.png" alt="Column Images" /></td>
<td><img src="image17.png" alt="Column Images" /></td>
<td><img src="image18.png" alt="Column Images" /></td>
</tr>
<tr>
<td>Column Aggregate</td>
<td><img src="image19.png" alt="Column Aggregate Images" /></td>
<td><img src="image20.png" alt="Column Aggregate Images" /></td>
<td><img src="image21.png" alt="Column Aggregate Images" /></td>
</tr>
<tr>
<td>Aggregate</td>
<td><img src="image22.png" alt="Aggregate Images" /></td>
<td><img src="image23.png" alt="Aggregate Images" /></td>
<td><img src="image24.png" alt="Aggregate Images" /></td>
</tr>
<tr>
<td>Irregular</td>
<td><img src="image25.png" alt="Irregular Images" /></td>
<td><img src="image26.png" alt="Irregular Images" /></td>
<td><img src="image27.png" alt="Irregular Images" /></td>
</tr>
</tbody>
</table>
3 Methodology

3.1 Network architecture

CNN is a type of Feed-forward Neural Network that incorporates convolutional computation and has a deep structure (LeCun et al. 2015; Gu et al. 2018). The key characteristics of CNN include weight sharing and local receptive field, which serve to reduce the number of network parameters. A typical CNN consists of various layer structures, including the input layer, hidden layer, and output layer. The hidden layer comprises the convolutional layer, pooling layer, fully connected layer, and more. As one of the representative algorithms of deep learning, CNN has been widely used in image classification tasks with its powerful data processing capability (Haq et al. 2023; Ahad et al. 2023). However, in practical applications, it is relatively difficult to obtain enough labelled image datasets for accurate classification. And training with substantial image datasets becomes challenging without access to high performance computer resources. Therefore, few people train new CNN models from scratch. In addition, most of the pretrained CNN are trained and optimised for convergence based on the ImageNet database, which has been achieved to minimise getting the false prediction error (Deng et al. 2009). Moreover, the pretrained image classification CNN has learnt a rich set of image features that can be used as a basis for learning new classification tasks. Therefore, to improve the accuracy of ice crystal habit classification, the most effective approach is to employ transfer learning. This entails using publicly available pretrained CNN model as a starting point and then training the CNN based on the ice crystal dataset to fine-tune the network's deeper parameters to adapt to the specific ice crystal shape classification task, ultimately achieving highly accurate ice crystal shape classification. In this paper, we use nine pretrained CNN models to perform transfer learning on the ice crystals dataset. These nine pretrained CNN models are as follows: TL-AlexNet, TL-Squeezenet, TL-Googlenet, TL-Efficientnetb0, TL-Mobilenetv2, TL-Resnet18, TL-Resnet50, TL-Resnet101 and TL-Shufflenet. These nine models are the classical and more commonly used in computer vision classification tasks. By inputting the ice crystal dataset into pretrained CNN model and iteratively training the ice crystals using the model's pretrained parameters, we ultimately obtain a CNN model suited for ice crystal shape classification.

3.2 Training details

The experimental operating platform used in this paper are MATLAB R2022A and Anaconda 3.4. The hardware device and system environment used is NVIDIA Quadro RTX 4000 with 64GB RAM and Windows OS.

The training details of the deep CNN model for ice crystal shape classification based on transfer learning are as follows. Each of the HVPS, 2D-S, and CPI probes contains 300 images for each ice crystal shape class. Firstly, we divided the ice crystal image dataset from these three probes into separate training set and test set in a ratio of 4:1, respectively. And then, the training and test sets for each of the three probes were combined together. Before training, we applied the same data preprocessing steps for each image in the ice crystal dataset: resizing all ice crystal particle images...
to a 227 × 227 and normalising the image pixel matrix. These operations to enhance model convergence and gradient descent speed. To prevent the network from overfitting and remembering specific features in the training images, we implemented data augmentation operations such as flip transformations, scaling, or random angle rotation, with a 50% probability before each input of particle images into the network. We selected the cross-entropy function as the loss function. To optimise the model parameters, the Sgdm (SGD with momentum) optimiser was used for stochastic gradient descent, and the decay rate of its momentum parameter (Sutskever et al., 2013) was set to 0.9. Additionally, the initial learning rate was set to η = 0.01, with the learning rate halved after every 10 epochs. The batch size was configured as 16, and the training epoch was set to 20. Subsequently, we loaded the pretrained CNN model and adjusted the output of the final fully connected layer to 9 to correspond to the nine ice crystal classes in our dataset. To obtain a suitable model for ice crystal habit classification, transfer learning was performed using the deep CNN models of TL-AlexNet, TL-Squeezenet, TL-Googlenet, TL-Efficientnetb0, TL-Mobilenetv2, TL-Resnet18, TL-Resnet50, TL-Resnet101, and TL-Shufflenet on the ice crystal dataset. Then, the results of different models during the training and testing processes were recorded, and the configuration of the model with the highest classification accuracy during testing is saved. Finally, the classification performance of the models is evaluated using the performance measure in section 3.3.

3.3 Performance measure

In machine learning, TP is a positive example correctly judged as a positive example; FP is a counterexample incorrectly judged as a positive example; TN is a counterexample correctly judged as a counterexample; FN is a positive example incorrectly judged as a counterexample. To evaluate the generalisation performance of the automatic image classifier more accurately and quantitatively, we use multiple performance measure to perform statistical analysis of model performance. The following is a description of the model performance measure:

(1) Accuracy

The fundamental measure for assessing a classification model's performance is accuracy, which represents the ratio of correctly classified samples to the total number of samples.

\[ A\_i (y, \bar{y}) = \frac{1}{N} \sum_{i=1}^{N} eq(y_i = \bar{y}_i) \]  \hspace{1cm} (1)

Where \(\bar{y}_i\) is the predicted label of the i-th image predicted by the model, \(y_i\) represents the corresponding true label, \(N\) is the total sample count, and \(eq(y_i = \bar{y}_i)\) is the value of 1 if the predicted label of the ice crystal image is the same as true label, and 0 otherwise.

(2) Cross-entropy loss

Cross-entropy loss reflects the difference between the true distribution of ice crystal labels and the predicted probability distribution. The formula is computed as (De Boer et al. 2005):

\[
-
\]

\[
-
\]
\[ H(p,q) = -\sum_{i=1}^{N} p(x_i) \log[q(x_i)] \]  \hspace{1cm} (2)

(3) Standard deviation

Standard deviation of the model recognition accuracy can be used to measure the robustness of the model and is calculated as:

\[ \sigma = \sqrt{\frac{\sum_{k=1}^{N} (A_{ck} - A_c)^2}{N - 1}} \]  \hspace{1cm} (3)

Where \( A_{ck} \) is the model recognition accuracy for the \( k \)-th category of ice crystal particle and \( A_c \) is the model recognition accuracy (\( k = 1,2,3...,9 \)).

(4) Precision

Precision is the number of ice crystal particles correctly identified as the \( k \)-th category as a proportion of the total number of ice crystals identified the \( k \)-th category.

\[ P = \frac{TP}{TP + FP} \]  \hspace{1cm} (4)

(5) Recall

Recall is the number of ice crystal particles correctly identified as the \( k \)-th category as a proportion of the total number of particles in \( k \)-th category.

\[ R = \frac{TP}{TP + FN} \]  \hspace{1cm} (5)

(6) F1-measure

F1-measure is the harmonic mean of precision and recall.

\[ F_1 = \frac{2TP}{2TP + FN + FP} \]  \hspace{1cm} (6)

(7) macro_Precision, macro_Recall and macro_F1

Precision, recall and F1-measure are calculated separately for each category of ice crystals and then averaged to obtain macro_Precision, macro_Recall and macro_F1. These three performance measures evaluate the overall performance of model. The calculation formula is as:

\[ macro_P = \frac{1}{N} \sum_{k=1}^{N} P_k \]  \hspace{1cm} (7)

\[ macro_R = \frac{1}{N} \sum_{k=1}^{N} R_k \]  \hspace{1cm} (8)
\[ marco_{F_i} = 2 \frac{marco_{P} \cdot marco_{R}}{marco_{P} + marco_{R}} \] (9)

4 Results and discussion

4.1 Experimental results of Ice crystal classification model

During the training and testing stages, we primarily employed accuracy and cross-entropy loss to evaluate and compare the 9 CNN models based on transfer learning. Subsequently, we selected the model with the highest accuracy on the test set to analyze other performance measures. Figure 2 displays the accuracy and cross-entropy loss values of the nine CNN models during training and testing. Training and testing were performed for 8100 iterations, and the results show that as the number of iterations increases, the accuracy of both the training set and test set increases rapidly, while the loss values decrease quickly for all CNN models. After about 15 epochs, all of them gradually become stable, with only slight fluctuations in change. For training, TL-Efficientnetb0 performs the least favorably, with the lowest accuracy and the highest cross-entropy loss value and TL-Shufflenet is also ineffective, with classification accuracy and cross-entropy loss only better than TL-Efficientnetb0. In contrast, TL-AlexNet achieves the best classification results with the highest accuracy and the lowest cross-entropy loss value. For testing stage, the TL-AlexNet also has the highest accuracy and the lowest cross-entropy loss value, while TL-Efficientnetb0, TL-Squeezenet and TL-Mobilenetv2 have relatively low classification accuracy and high cross-entropy loss, and the rest of the classification models are next to TL-AlexNet. The difference in testing accuracy between the worst and best performing CNN models is about 7% (see the top right panel of Figure 2). Overall, the TL-AlexNet classification model exhibits the best classification performance on ice crystal dataset in this paper.

To better distinguish the effectiveness of classification models, we compared the highest classification accuracy of each CNN model based on transfer learning on the test set with corresponding standard deviation and the training time of model (see Table 4). The highest accuracy of TL-AlexNet and TL-Resnet50 models on the test set are over 97%, while TL-AlexNet has the best accuracy with 97.16%. In contrast, the highest classification accuracy of TL-Squeezenet and TL-Efficientnetb0 is significantly lower than other ice crystal classification transfer learning models, both less than 95%. In addition, the best classification accuracy rates of TL-Googlenet (96.17%), TL-Mobilenetv2 (95.31%), TL-Resnet18 (95.93%), TL-Resnet101 (96.23%), and TL-Shufflenet (95.80%) are lower than TL-AlexNet. And standard deviation of the recognition accuracy of the TL-AlexNet model is smaller than that of the other seven transfer learning models except TL-Resnet50, indicating TL-AlexNet shows better robustness. Additionally, as shown in Table 4, the training times for TL-AlexNet, TL-Squeezenet, and TL-Resnet18 are all within 3000 s. Among them, the minimum training time is 1799 s for the TL-AlexNet, while the maximum training time for TL-Efficientnetb0 reaches 13,607 s. In summary, TL-AlexNet, with the
stability, generalization ability, and exceptional accuracy, seems to be the best choice for ice crystal habit classification.

![Fig. 2. Accuracy (top) and cross-loss entropy curves (bottom) of nine CNN models based on transfer learning during training and testing stages.](image-url)

### Table 4

<table>
<thead>
<tr>
<th>Transfer learning model</th>
<th>$A_c$ (%)</th>
<th>$\sigma$</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TL-AlexNet</td>
<td>97.16</td>
<td>0.0214</td>
<td>1799</td>
</tr>
<tr>
<td>TL-Squeezenet</td>
<td>94.07</td>
<td>0.0428</td>
<td>2017</td>
</tr>
<tr>
<td>TL-Efficientnetb0</td>
<td>94.01</td>
<td>0.0450</td>
<td>13607</td>
</tr>
<tr>
<td>TL-Googlenet</td>
<td>96.17</td>
<td>0.0255</td>
<td>3240</td>
</tr>
<tr>
<td>TL-Mobilenetv2</td>
<td>95.31</td>
<td>0.0327</td>
<td>4341</td>
</tr>
<tr>
<td>TL-Resnet18</td>
<td>95.93</td>
<td>0.0290</td>
<td>2148</td>
</tr>
<tr>
<td>TL-Resnet50</td>
<td>97.10</td>
<td>0.0198</td>
<td>5517</td>
</tr>
<tr>
<td>TL-Resnet101</td>
<td>96.23</td>
<td>0.0270</td>
<td>12693</td>
</tr>
<tr>
<td>TL-Shufflenet</td>
<td>95.80</td>
<td>0.0240</td>
<td>4507</td>
</tr>
</tbody>
</table>

### 4.2 Evaluation of TL-AlexNet Model

To provide a comprehensive assessment of TL-AlexNet's performance, we conducted a detailed evaluation of precision, recall, and F1-measure for each class of ice crystal.
particle classified by the TL-AlexNet model (see Figure 3). Figure 3 shows that the TL-AlexNet model achieves a 100% recall for the sphere, while the recall for the aggregate is relatively low at 92.22%. Column and sphere are classified most effectively, likely owing to their unique and easily distinguishable characteristics. In contrast, aggregate is the worst classified, probably due to the fact that the formation of aggregate has some structural similarity in the basic morphology of other ice crystal categories. Additionally, the precision, recall, and F1-measure for the classification of particles exceeded 95% for all categories except for the recall and F1-measure of aggregate. This indicates that the TL-AlexNet classification model shows better classification results in every category of ice crystals except the aggregates category, which has a wide range of classification effects. And macro_Precision, macro_Recall and macro_F1 of nine ice crystal categories, which are all higher than 97%. This demonstrates the stable and efficient overall classification effectiveness of the TL-AlexNet model, which is an automated and highly accurate ice crystal shape classification model.

Next, we further assessed the classification ability of the TL-AlexNet model on the test set using a confusion matrix. Confusion matrix is one of the most common methods in judging the effectiveness of classification models, which is not only used to judge the performance of classification models, but also a basic, intuitive and computationally simple model performance measure. It assesses a classification model by comparing predictions to actual categories. The diagonal elements in the confusion matrix correspond to the samples that were accurately classified. Figure 4 presents the confusion matrix results for the TL-AlexNet model when classifying ice crystal on the test set. The results show that most ice crystal shape categories are effectively distinguished. Specifically, the TL-AlexNet model exhibits minimal prediction errors in the classification of irregular, sphere, dendritic, column aggregate, and column categories. However, some problems arise when distinguishing between aggregate and plate categories, possibly due to the infrastructure of the aggregate.
contain plate crystal structure. Nevertheless, in terms of the overall classification effect of the test set, the classification error of TL-AlexNet is still low. In summary, TL-AlexNet shows high accuracy in habit classification of ice crystal data from CPI, 2D-S and HVPS, which meets the requirements of the research.

![Confusion matrix of ice crystal particle classification results for the TL-AlexNet classification model on the test set.](image)

**Fig. 4.** Confusion matrix of ice crystal particle classification results for the TL-AlexNet classification model on the test set.

We also investigated the ability of the TL-AlexNet model on the three test sets from the CPI, 2D-S and HVPS probes. Figure 5 displays the confusion matrix for the TL-AlexNet on these three probe test sets. Overall, sphere morphology performs best in the ice crystal dataset for all three probes, while aggregate is the worst classified in the dataset for all three probes. Figure 5 reveals that the TL-AlexNet model has minimal prediction errors in classifying irregular, sphere, plate, dendritic, column, and rosette categories for the CPI test set, sphere, dendritic, column aggregate, and graupel categories for the 2D-S test set, and irregular, sphere, dendritic, column, and column aggregate categories for the HVPS test set, with only 0 to 1 sample prediction errors. However, for 2D-S test set, the overall classification ability of the TL-AlexNet model was slightly lower compared to CPI and HVPS test sets.

We evaluated the macro average values of precision, recall, and F1 measures for the test sets of three different probes (refer to Table 5). As indicated in the table, the TL-AlexNet classification model outperforms other two probe test sets in all evaluation measure on the CPI probe test set, and the evaluation measure are higher than 0.97 on both the CPI and HVPS probe test sets. Although the resolution of 2D-S probe is higher than HVPS probe, the TL-AlexNet classification model is more effective in classifying the HVPS probe dataset in the ice crystal classification task. The reason may be that higher resolution ice crystal images may contain more details, which may increase the difficulty of feature extraction and learning. If the model fails to take full advantage of these details, it may lead to poorer classification results. Moreover, the vast majority of the shape image data for HVPS probe, except for the
data of the column aggregate, come from the OLYMPEX field observation project. The relatively homogeneous geographic location of the HVPS image data makes it easier for the model to generalise to the ice crystals from this project.

Fig. 5. Confusion matrix of ice crystal particle classification results for the TL-AlexNet classification model on the CPI (a), 2D-S (b) and HPVS (c) probe test sets.

Table 5. Font sizes of headings. Table captions should always be positioned above the tables.

<table>
<thead>
<tr>
<th></th>
<th>CPI</th>
<th>2D-S</th>
<th>HPVS</th>
</tr>
</thead>
<tbody>
<tr>
<td>marco_P</td>
<td>0.9798</td>
<td>0.9596</td>
<td>0.9770</td>
</tr>
<tr>
<td>marco_R</td>
<td>0.9796</td>
<td>0.9593</td>
<td>0.9759</td>
</tr>
<tr>
<td>marco_F</td>
<td>0.9797</td>
<td>0.9594</td>
<td>0.9765</td>
</tr>
</tbody>
</table>

5 Conclusion and outlook

We manually constructed an ice crystal dataset comprising nine habit categories, totaling 8100 images, utilizing images observed by airborne CPI, 2D-S, and HVPS probes. By the transfer learning method, we utilized 9 pretrained CNN models to establish automated classification models for ice crystal shapes. We compared the classification accuracy, cross-loss entropy, standard deviation, and training time of
each model during training and testing to compare the model performance, and ultimately selected the TL-AlexNet model with the best classification accuracy during testing. For a comprehensive assessment of the TL-AlexNet model's performance across nine ice crystal shape categories, we employed multiple performance measures to gauge its classification effectiveness on the ice crystal image test set. The results of this evaluation demonstrated that the TL-AlexNet model achieves a classification accuracy of 97.16%. Additionally, the classification model confused a small proportion of ice crystals, for example, misclassifying aggregate as column aggregate or plate, and misclassifying rosette as dendritic, but the confusion errors are only within 4%.

Furthermore, we analysed the classification ability of TL-AlexNet model on the ice crystal test sets of CPI, 2D-S and HVPS, respectively. Despite the fact that HVPS images have a low pixel count and are in binary, the TL-AlexNet model is still able to efficiently learn features related to the defined ice crystal shape categories. Based on the results and analyses of the above experiments, the TL-AlexNet classification model proposed in this study not only outperforms the traditional ice crystal classification methods, but also can effectively and automatically classify ice crystals of different shape categories from two types of imagers.

However, it's worth noting that the current accuracy of the TL-AlexNet model for recognizing aggregate, rosette, and plate shapes still requires improvement. Future research should further extend the ice crystal dataset and improve the model classification performance so that it can distinguish different ice crystal shapes more accurately. In addition, images captured by probes such as PIP and four-level greyscale CIP, or images taken by other types of imagers such as those using optical sensing techniques, could be considered for addition to the ice crystal dataset for use in the development of more powerful models for automatic classification of ice crystal shapes.

Funding

This work is supported by Instrument and Equipment Function Development Project of Chinese Academy of Sciences (E3661716), Technical Support Talent of Chinese Academy of Sciences in 2022 (E3661416), National Natural Science Foundation of China (41775166 and 41705142), Beijing Information Science & Technology University Project (20190091).

References


Lohmeier SP, Sekelsky SM, Firda JM, Sadowy GA, McIntosh RE (1997) Classification of particles in stratiform clouds using the 33 and 95 GHz polarimetric cloud profiling radar system (CPRS). IEEE transactions on geoscience and remote sensing 35(2): 256-270. https://doi.org/10.1109/36.563264


Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Highlights.pdf