

## Article

# Cloud-MobiNet: An Abridged Mobile-Net Convolutional Neural Network Model for Ground-Based Cloud Classification

Emmanuel Kwabena Gyasi  and Purushotham Swarnalatha \*

School of Computer Science and Engineering, VIT University, Vellore 632014, India

\* Correspondence: pswarnalatha@vit.ac.in; Tel.: +91-9443630735

**Abstract:** More than 60 percent of the global surface is covered by clouds, and they play a vital role in the hydrological circle, climate change, and radiation budgets by modifying shortwaves and longwave. Weather forecast reports are critical to areas such as air and sea transport, energy, agriculture, and the environment. The time has come for artificial intelligence-powered devices to take the place of the current method by which decision-making experts determine cloud types. Convolutional neural network models (CNNs) are starting to be utilized for identifying the types of clouds that are caused by meteorological occurrences. This study uses the publicly available Cirrus Cumulus Stratus Nimbus (CCSN) dataset, which consists of 2543 ground-based cloud images altogether. We propose a model called Cloud-MobiNet for the classification of ground-based clouds. The model is an abridged convolutional neural network based on MobileNet. The architecture of Cloud-MobiNet is divided into two blocks, namely the MobileNet building block and the support MobileNet block (SM block). The MobileNet building block consists of the weights of the depthwise separable convolutions and pointwise separable convolutions of the MobileNet model. The SM block is made up of three dense network layers for feature extraction. This makes the Cloud-MobiNet model very lightweight to be implemented on a smartphone. An overall accuracy success of 97.45% was obtained for the CCSN dataset used for cloud-type classification. Cloud-MobiNet promises to be a significant model in the short term, since automated ground-based cloud classification is anticipated to be a preferred means of cloud observation, not only in meteorological analysis and forecasting but also in the aeronautical and aviation industries.



**Citation:** Gyasi, E.K.; Swarnalatha, P. Cloud-MobiNet: An Abridged Mobile-Net Convolutional Neural Network Model for Ground-Based Cloud Classification. *Atmosphere* **2023**, *14*, 280. <https://doi.org/10.3390/atmos14020280>

Academic Editor: Stephan Havemann

Received: 6 January 2023

Revised: 27 January 2023

Accepted: 30 January 2023

Published: 31 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Cloud-MobiNet; remote sensing; artificial intelligence; convolutional neural network; climate change; machine learning; radiation budget; weather forecast; cloud prediction; smartphone

## 1. Introduction

Cloud observation forms part of the core duties of meteorologists as it helps them to make informed decisions in weather forecasting. Clouds are masses of water droplets or ice particles floating in the sky. The type of cloud, the total amount of cloud (cloud cover), and the height of the cloud are the three main factors that meteorologists consider while observing clouds. Essentially, there are ten main types of clouds. They are classified as high clouds, medium clouds, or low clouds by the World Meteorological Organization (WMO). High clouds include cirrus (Ci), cirrocumulus (Cc), and cirrostratus (Cs), whereas medium clouds include altocumulus (Ac), altostratus (As), and nimbostratus (Ns). Stratus (St), stratocumulus (Sc), cumulus (Cu), and cumulonimbus (Cb) are some of the low clouds.

More than 60 percent of the global surface is covered by clouds, and they play a vital role in the hydrological circle, climate change, and radiation budgets by modifying shortwaves and longwave [1]. Each cloud, as well as the combination of two or more clouds, has its meteorological phenomena. For instance, in the aviation industry, though aircraft can experience turbulence in clear air, clouds contribute to the greater amount of turbulence experienced by the aircraft. Although most commercial airlines fly above much of the clouds, they still fly through clouds during landing and takeoff at the airports. A

typical commercial jet has a cruising altitude of around six to seven miles (nine to 11 km) above sea level; hence, on a long-distance flight, a plane will generally be above most clouds, except for cirrus and the towering cumulonimbus.

Air passengers experience a bumpy flight when the aircraft enters clouds ranging from puffy cumulus clouds, which are also known as fair-weather clouds, to monstrous cumulonimbus clouds with their characteristic anvil-shaped tops, billowing sides, and ominously dark bases. When clouds are cooler than the surrounding air, the contrast in density between the clouds and the surrounding air creates a sort of “pothole” in the sky, resulting in a less-smooth flight. Storm clouds, such as cumulonimbus, are the types of clouds that most pilots want to avoid. Cumulonimbus clouds generally contain heavy rain, lightning, hail, strong winds, and occasionally tornadoes. Pilots and air traffic control pay close attention to the weather and route flights around these types of storm clouds. Similarly, clouds can block light and heat from the Sun, making Earth’s temperature cooler.

In meteorology, cloud cover is measured in oktas, or eighths of the sky. The observers divide the sky into eight boxes in their minds as they look up at it, and then picture all the clouds they can see being crammed into these boxes. Then, they count how many boxes the cloud fills. The number of boxes they will receive corresponds to the number of oktas of clouds. Zero oktas represent a complete absence of cloud, 1 okta represents a cloud amount of one-eighth or less, but not zero, 7 oktas represent a cloud amount of seven-eighths or more, but not full cloud cover, while 8 oktas represent full cloud cover with no breaks.

Meteorologists also use terminology to convey generally how cloudy it is; for instance, “few clouds” refers to 1 to 2 oktas, the “scattered cloud” refers to 3 to 4 oktas, where about half the sky is covered, and the “broken cloud” is 5 to 7 oktas, where much of the sky is covered and “overcast” is 8 oktas of cloud with no breaks in the cloud. When there is no presence of clouds in the sky, the reported cloud amount as “nil”.

Similarly, at most synoptic observing stations, cloud base (height) is readily measured by an instrument to a reasonable level of accuracy. The cloud base recorder employs a pulsed diode laser LIDAR (light detection and ranging) technology [2], whereby short laser pulses (eye safe) are sent out in a vertical or near vertical direction. The backscatter caused by reflection from the surface of the cloud, precipitation, or other particles is analyzed to determine the height of the cloud base. Many modern cloud base recorders are capable of detecting up to three cloud layers simultaneously.

However, some meteorologists at weather observation stations still rely heavily on the manual estimation of clouds’ height. The observers determine the types of clouds present in the sky and use a generally estimated range of each of the 10 types of clouds’ heights to estimate the height of the cloud observed. For example, low clouds, which include cumulus clouds, can form anywhere from near the surface up to 2000 m (6500 feet). Middle clouds form at altitudes of 2000–4000 m (6500–13,000 ft) above ground near the poles, 2000 to 7000 m (6500–23,000 ft) at mid-latitudes, and 2000 to 2600 m (6500–25,000 ft) at the tropics. High clouds have base heights of 3000 to 7600 m (10,000–25,000 ft) in polar regions, 5000 to 12,200 m (16,500–40,000 ft) in temperate regions, and 6100 to 18,300 m (20,000–60,000 ft) in the tropical regions.

The most important aspect of cloud observation is cloud identification, as opposed to cloud base (height) and cloud amount (cloud cover). It is quite challenging for meteorologists to determine the type of cloud present at any particular time due to the clouds’ similarity in shape, color, form, and texture. It depends on the observer’s knowledge, experience, and color vision to identify clouds accurately.

Weather forecast reports are critical to areas such as transportation (air, sea, and land), agriculture, energy, the environment, and the general public. Given the potential for misclassification, clouds often cause weather forecasts to be inaccurate, thereby exposing lives and property to extreme weather disasters.

To forestall such occurrences, a variety of measurement equipment, including satellite-based and ground-based remote sensors, is used to collect the necessary cloud data for classification tasks. Ceilometers are standalone devices that use laser-based light detection

and range (LIDAR) technology to measure the height profiles of clouds and aerosols [2]. However, there are difficulties involved in using ceilometers to gauge cloud height. They do not display the type of cloud observed, are extremely expensive and, when used as a standalone instrument, are large and heavy, difficult to use, and require specialist knowledge. Furthermore, they are fragile and unreliable due to their susceptibility to weather conditions. Examples of ceilometers are the Eliason CBME80B and the Vaisala CL31. Additionally, clouds may be observed, from a downward perspective, by satellite-based weather equipment across wide regions. However, their spatial resolution is too low to represent small-scale cloud features across large regions [3].

Several types of studies have been undertaken in an attempt to find a better method of cloud observation. Although the acquired accuracy was not sufficient owing to dataset constraints and the approach utilized, Zhuo et al. [4] developed a color-census transformation to extract texture and structural information from a color-sky image. Refs. [4–6] have advocated using hand-crafted elements, including color, texture, structure, and others, to classify ground-based cloud images in their investigations. He et al. and Labati et al. [7,8] also explored the use of machine-learning-based convolutional neural networks to classify ground-based clouds. Furthermore, Shi et al. [9], for example, denoted each ground-based cloud image with deep convolutional activation-based features acquired through the pooling of the convolutional activations of each feature map, and ran tests on two datasets with 784 cloud images in five classes and 1500 images in seven classes.

In another experiment to classify 1231 cloud images into nine groups, Ye et al. [10] retrieved features from various convolutional layers, in which discriminative-local patterns are picked and subsequently represented via the Fisher vector. The learning-group-patterns method extracts cloud properties using wireless sensor networks [11]. The combined accuracy of the two databases, on the other hand, is 81%. The task-based graph convolutional network (TGCN) model, which combines graph computation with a deep network to classify ground-based clouds, was also suggested by [12]. CNNs, which are a type of deep-learning architecture, have previously excelled in a variety of fields, including pattern recognition and computer vision [13,14].

Orthodox methods fail to describe and extract the characteristics of clouds because of the convolution of cloud texture and patterns, but CNNs can learn increasingly intricate patterns and discriminative textures from vast pre-trained and labeled datasets [15]. Additionally, convolutional neural networks typically feature tiered characteristics-extraction frameworks. On the whole, CNNs shallow layers capture fine textures, such as edge and shape, but the deeper layers represent high-level semantic data based primarily on pixels. Previous research findings indicate that both semantic features and textures are vital for cloud characterization [4].

Despite CNN's image-classification prowess, few researchers have evaluated its accuracy and effectiveness in cloud classification. Cloud images have different qualities in cloud representation, particularly for contrails, since they are distinctive texture images with odd forms. Given the incredibly varied and difficult-to-classify characteristics of clouds, the adaptive-learning aspect of neural network classification provides a high-accuracy and computationally efficient alternative to cloud classification [16]. Fabel et al. [17] used about 300,000 semantic segmentations of ground-based all-sky images (ASIs) in two different pretext tasks for pretraining. One of them pursues an image reconstruction approach, while the other is based on the deep cluster model, an iterative procedure of clustering and classifying the neural network output. Li et al. [18] classified ground-based cloud images by using contrasting self-supervised learning to pre-train the deep model with a contrastive loss and momentum update-based optimization. Liu et al. [12] classified ground-based cloud images by using a transformer-based GCI classification method that combines the advantages of the CNN and transformer models. Toğaçar et al. [19] also classified clouds by using super-resolution, semantic segmentation approaches, and binary sailfish optimization methods with deep learning models.

From cloud representation to cloud classification, the spatial format of fully connected (FC) features, as well as the local texture data of shallower convolutional layers, are crucial. Although these studies helped with cloud classification, a reliable cloud classification is not yet possible. As a result, work on an automated method that can accurately classify ground-based cloud images is continuing.

Therefore, we present Cloud-MobiNet, a robust CNN model for ground-based cloud classification that not only provides excellent classification accuracy but is also compact, efficient, portable, and can be used on smartphones. This study focuses on finding a more reliable way of classifying clouds in real-time to curtail the problems associated with employing ceilometers and other traditional methods, such as human cloud observation, which depends on the observer's training, expertise, and color vision.

The different sections of this study are arranged as follows. The Cloud-MobiNet model is explained, followed by descriptions of the dataset and data preprocessing utilized in the experiment, which is followed in turn by an explanation of Section 2's model-training procedure. The experimental results, model-performance assessment, classification report, confusion matrix, and model deployment on a smartphone are then covered in depth in Section 3. The results are discussed in Section 4. Finally, Section 5 provides a summary of the findings.

## 2. Materials and Methods

### 2.1. Model Architecture

Cloud-MobiNet is an abridged convolutional neural network model, based on MobileNet, used for classifying ground-based cloud images. The architecture of Cloud-MobiNet is divided into two blocks, namely the MobileNet building block and the Support MobileNet block (SM block). The MobileNet building block consists of the weights of the depthwise separable convolutions and pointwise separable convolutions of the MobileNet model, while the SM block is made up of three dense network layers of neurons. The three dense layers are purposely introduced for the extraction of features from the given cloud images.

The motive behind this methodology is to utilize the transfer of knowledge from the source domain  $D_S$  of natural images to our target domain of ground-based cloud images  $D_T$ .

Therefore, a predicted target image  $y_{Ti}$  is calculated as follows:

$$y_{Ti} = D_S + D_T, \quad (1)$$

where the predicted target image  $y_{Ti}$  is the 11 categories of class images [ $y_{Ti} = \{Ci, Cc, Cs, Ac, As, Ns, St, Sc, Cu, Cb, Ct\}$ ].

The target domain  $D_T$  is the total number of dense layers in the (SM Block)

$$[D_T = \{D_{L1} + D_{L2} + D_{L3}\}]$$

This means that a target classifier must be trained such that, given a ground-based cloud image from the target domain  $D_T$ , we obtain a prediction about the classes of the image in 11 categories. Therefore, a predicted target image  $y_{Ti}$  is obtained through the transfer of the weights of the source domain  $D_S$  and the weight of the extracted features of the target domain's image by the 3 dense layers  $D_T$ . This is specified in Equation (1). To reduce the weight of the MobileNet parameters, the top 1000 neurons of MobileNet network layers were frozen. The supplied image measures  $224 \times 224 \times 3$ , which represents the width, height, and RGB channel, respectively. MobileNet is meant to be utilized in mobile applications [20], since it is a lightweight deep neural network. This makes Cloud-MobiNet very light, and with a significantly reduced number of parameters compared to the already existing cloud-classification architectures, such as CloudNet [21,22]. Figure 1 shows the Cloud-MobiNet architecture.

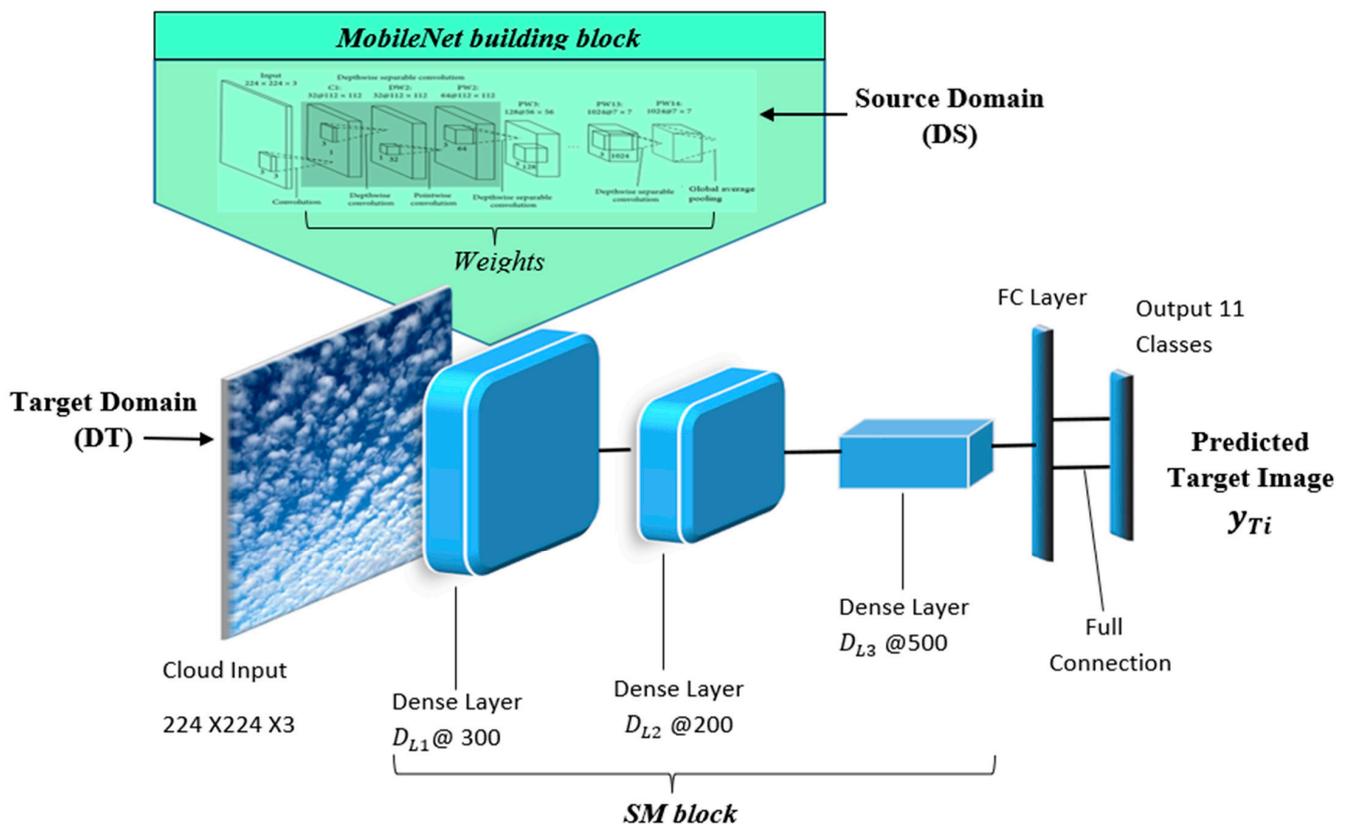


Figure 1. Cloud-MobiNet architecture.

Figure 2 is the diagram of the MobileNet building block inside Cloud-Mobinet architecture.

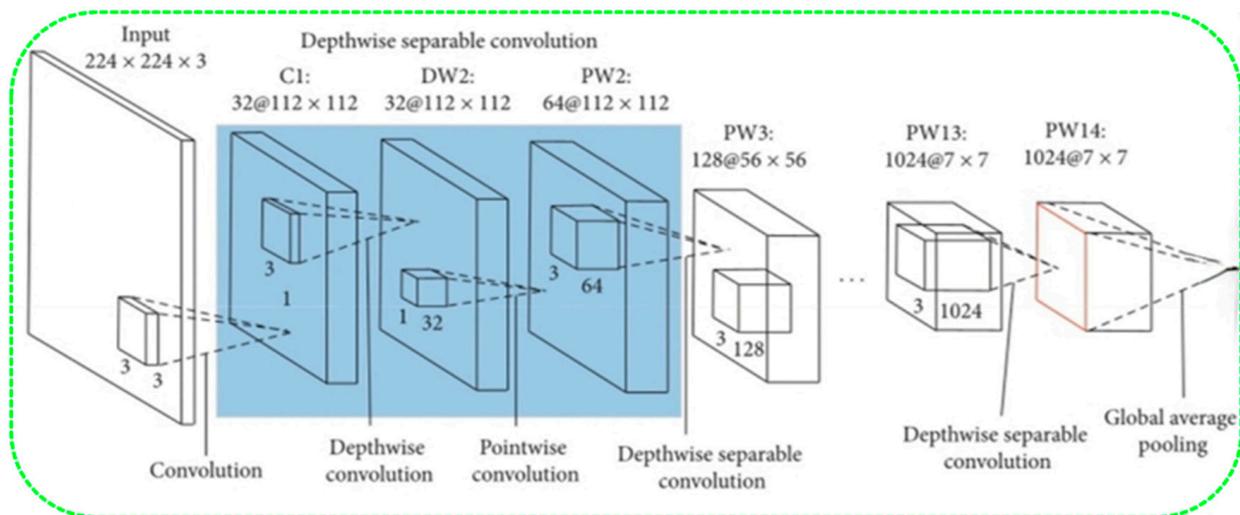
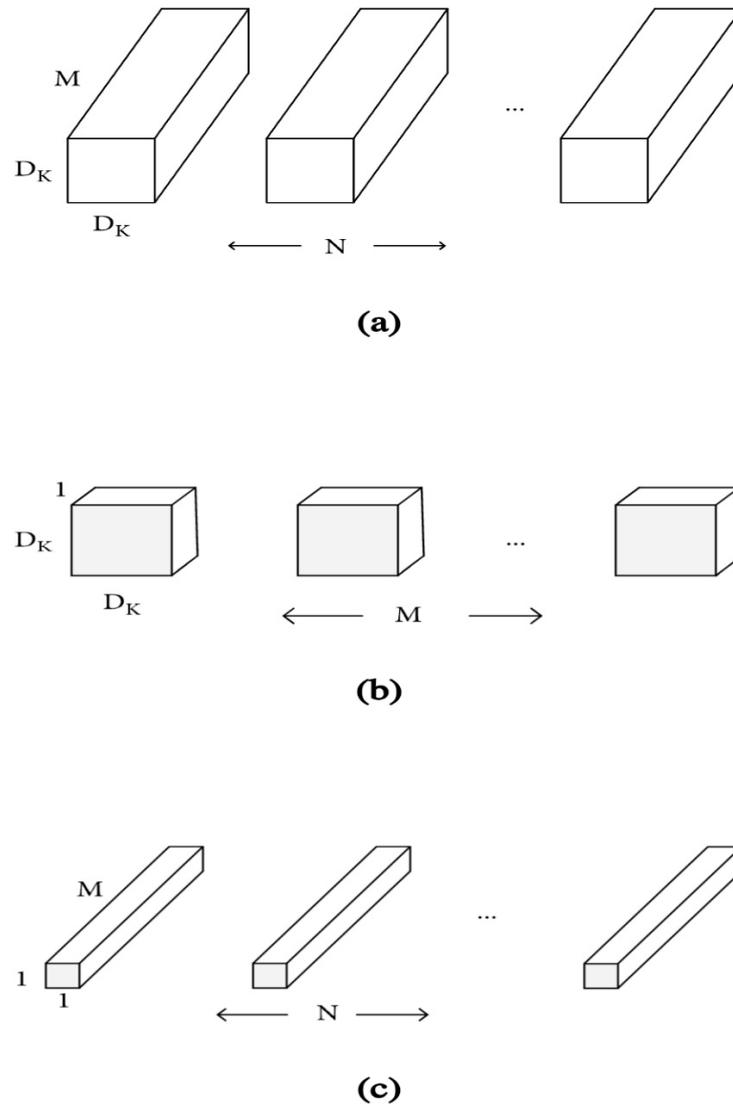


Figure 2. MobileNet building block.

MobileNet networks’ depthwise separable convolutions minimize the size and complexity of the model. Since the models are small, they have fewer model parameters. The depthwise separable convolution operation approximates a standard convolution using a form of factorized convolution. The approximation operation factorizes a standard convolution into a depthwise convolution, and a  $1 \times 1$  convolution called the pointwise convolution. A  $1 \times 1$  convolution is then applied to the outputs of the depthwise and pointwise convolutions to combine them. Using a conventional convolution, inputs are

combined and filtered into a new set of outputs in one step. The depthwise separable convolution divides this into two layers, one for filtering and one for combining [20]. The result of this factorization is a significant decrease in computation and model size and is essential for embedded and mobile device deployments. Figure 3 shows how (a) a standard convolution is factorized into (b) a depthwise convolution and (c) a  $1 \times 1$  pointwise convolution.



**Figure 3.** Standard convolutional filters and depthwise separable filters. (a) Standard convolutional filters. (b) Depthwise convolutional filters. (c) Point convolutional filters.

### 2.2. The Data

The two options for continuously observing clouds are ground-based cloud observation and satellite-based cloud observation. As a result, the majority of cloud classification uses ground-based or satellite-based cloud images. A ground-based Cirrus Cumulus Stratus Nimbus (CCSN) cloud dataset, which is well-known to many academics, was employed in this work. The CCSN dataset’s quality is undeniable, and it also includes 11 different types of clouds. Generally, 10 different types of clouds have been identified by the WMO. In addition to these 10 unique cloud types, there is another type of artificial cloud, called the contrail, which is created by high-flying jet planes, formed of water droplets evaporated from water vapor in jet-engine exhaust. The CCSN dataset is owned by Harvard University and is openly accessible to the public. The dataset contains 2543 cloud images prepared

by experienced meteorological professionals following the WMO genera-based classification recommendation [21]. The images are in JPEG format with a fixed resolution of  $256 \times 256$  pixels, large lighting variations, and in-class variations. Table 1 shows the description of the cloud types in the CCSN cloud dataset. The initial quantity column represents the number of cloud images before performing the augmentation, while the final quantity column denotes the number of images after the augmentation process.

**Table 1.** Details of CCSN cloud dataset.

Cloud Name	Symbol	Initial Quantity	Final Quantity	Characteristics
Cirrus	Ci	139	2665	Feathery, wispy clouds of ice crystals
Cirrostratus	Cs	287	4773	Ice crystals, milky, translucent veil cloud
Cirrocumulus	Cc	268	4623	White flakes, fleecy clouds forming ripples
Alto cumulus	Ac	221	3924	White or gray with shading and rounded clumps
Altostratus	As	188	3424	Mainly gray or blueish clouds, opaque and ice crystals
Cumulus	Cu	182	3371	Cauliflower shape, fluffy, occasioned by rain or snow showers
Cumulonimbus	Cb	242	4274	Thunder cloud, icy, anvil-shaped top noted by heavy rain
Nimbostratus	Ns	274	4610	Rain cloud, grey cloud with a dark and a vague outline
Stratocumulus	Sc	340	5520	Compound dark grey layer cloud rollers or banks
Stratus	St	202	3625	Low cloud, causes fog, drizzle, or fine precipitation
Contrails	Ct	200	3625	Aircraft engine exhausts generate these line-shaped clouds
	Total	2543	44,434	

### 2.3. Preparation of Dataset

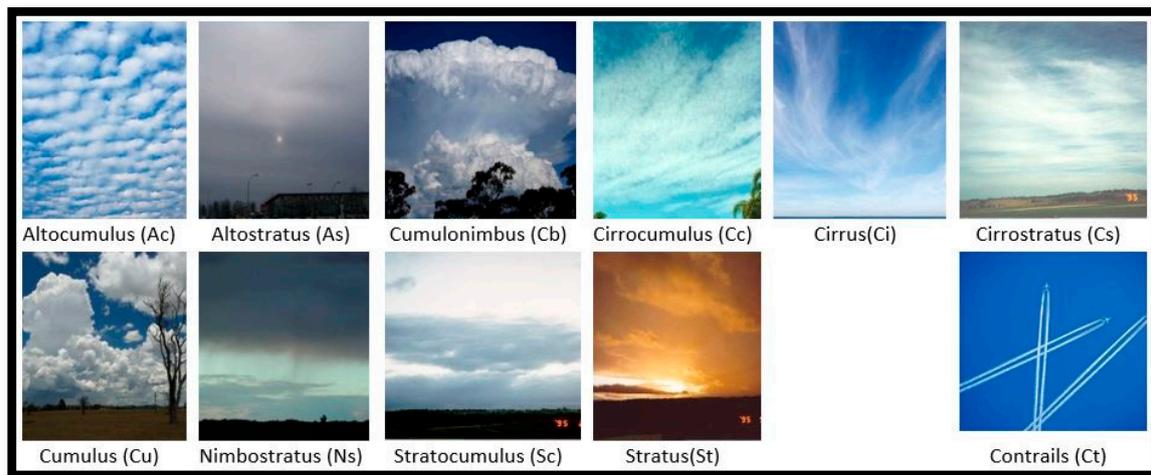
Neural network models thrive on the availability of huge and expertly labeled images. Image-data augmentation is a method that can be employed to substantially increase the amount of data artificially, through a modified version of images in the dataset. Several kinds of research have attested to the benefits of using a large dataset for the training of deep-learning neural network models and how an increased dataset makes models very skillful and robust. Through the use of augmentation techniques, several variations of images can be generated, which increases the fitted models' capacity to transfer what they have learned to new images.

During this experiment, we used Keras' deep-learning neural network library, which includes the ImageDataGenerator class, which allowed us to fit models using image-data augmentation. This ImageDataGenerator class supports a variety of techniques and pixel-scaling methods. We used image zoom, image shifts, and image flips, as well as image rotations, such as the *zoom\_range* argument, *width\_shift\_range*, and *height\_shift\_range* arguments, *horizontal\_flip* argument, and *rotation\_range* argument, respectively. The counter was set to 20, and this generated an additional 41,891 distinct images, which resulted in a very substantial cloud-image dataset for our experiment.

The CCSN data contain  $256 \times 256$  cloud images with 3 color channels and 11 classes. The training set, validation set, and test set are the three subsets of the data. The validation set forms 30% of the total data and the training set forms 70% of the total data. Finally, 110 randomly selected RGB cloud images from each of the 11 classes make up the test set. Figure 4 shows some samples of cloud images from the CCSN dataset.

### 2.4. Model Training

The Cloud-MobiNet model is trained on a Jupyter notebook hosted by Anaconda Integrated Development Environment (IDE) and runs on a 64-bit operating system, Intel(R) Core (TM) i7—7500U CPU @ 2.70 GHz 2.90 GHz processor, 12.0 GB RAM, and Windows 10 Education. Furthermore, the model was trained in TensorFlow [23] with Keras backend using Adam as our optimization method and a learning rate of 0.0001, batch size of 22, and rescaling our image to  $224 \times 224$ .

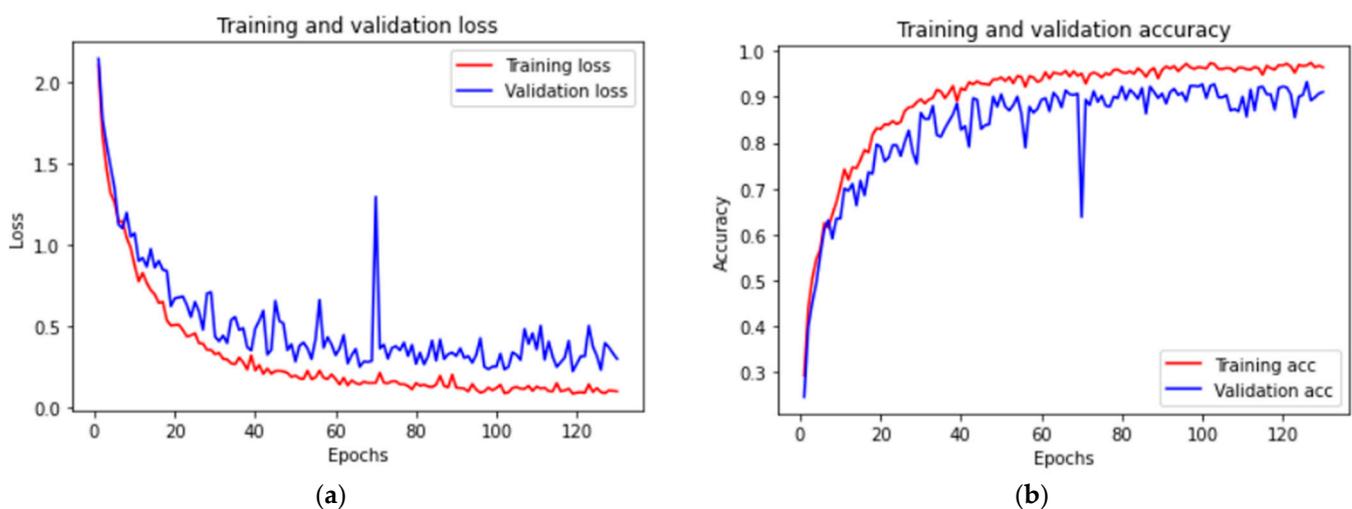


**Figure 4.** Cloud samples of the CCSN dataset.

Some cloud features are not distinct, which makes it extremely difficult to distinguish one from another. As a result, extracting features from them by directly applying FC layers and convolutional layers does not yield high accuracy [9,10]. To curtail this drawback, we introduced 3 dense layers without a dropout layer in our Cloud-MobiNet. For the model to better fit the training data with each iteration, we initially introduced early stopping as a form of regularization. This gave us an idea of the number of epochs we should run to avoid over-fitting and overtraining the model. Through this process, we arrived at running 130 epochs. The running time of the model was 4 h and 38 min with an average epoch time of 127 s and an average memory utilization of 3.5 GB.

### 3. Results

After 4 h and 45 min of training and validation of our Cloud-MobiNet model, the model gave an accuracy of 97.45% and a loss of 0.07624, representing 0.76% of the total. Figure 5a,b shows a graph of training and validation loss, and training and validation accuracy, respectively.



**Figure 5.** (a,b) The graph of training and validation loss, and training and validation accuracy, respectively.

As shown in Supplementary Document: SD1\_Cloud-MobiNet\_Codes, we carried out testing of the model to assess its performance by using the test dataset. Table 2 shows the classification report of the classes predicted by the Cloud-MobiNet model. The classification report was auto-generated as we employed the machine-learning library Scikit-Learn's

simple syntax: “from sklearn.metrics import classification\_report”. Nevertheless, we explain the mathematical principles behind this classification report based on four methods to determine whether the extrapolations are correct or incorrect.

**Table 2.** Classification report.

Cloud	Precision	Recall	F1-Score
Ac	1.00	1.00	1.00
Cb	1.00	1.00	1.00
As	1.00	0.90	0.95
Cc	1.00	0.90	0.95
Ci	0.91	1.00	0.95
Cs	0.91	1.00	0.95
Ct	1.00	1.00	1.00
Cu	1.00	1.00	1.00
Ns	1.00	0.80	0.89
Sc	1.00	1.00	1.00
St	0.83	1.00	0.91
<b>Accuracy</b>			<b>0.96</b>
<b>Macro Average</b>	<b>0.97</b>	<b>0.96</b>	<b>0.96</b>
<b>Weighted Average</b>	<b>0.97</b>	<b>0.96</b>	<b>0.96</b>

True Negative (*tn*): Indicates the case was and was anticipated to be negative. True-Positive (*tp*): Denotes an instance that was both positive and projected to be positive. False-Negative (*fn*): This means the case was positive, yet the aftereffect was predicted to be negative. False-Positive (*fp*): This means the case was negative, but the outcome was predicted to be positive.

The precision metric indicates how accurate our predictions were to the mark. The Cloud-MobiNet model can detect a negative instance as positive. For each class, it is determined as “the ratio of true positives to the sum of true positives and false positives”. Precision simply refers to the accuracy of positive predictions, which is calculated as follows:

$$Precision = \frac{tp}{(tp + fp)} \tag{2}$$

A recall is the ability of the Cloud-MobiNet model to discover all positive occurrences. A recall is calculated as follows:

$$Recall = \frac{tp}{tp + fn} \tag{3}$$

The F1 score is a weighted harmonic mean of precision and recall, with 1.0 being the greatest and 0 representing the poorest. The following is how the F1 score is determined:

$$F1 - Score(\beta) = \frac{(1 + \beta^2)tp}{(1 + \beta^2)tp + \beta^2fp + fn} \tag{4}$$

where  $\beta$  is taken to be 1.

Accuracy is calculated as:

$$Accuracy = \frac{\sum_c tp_c}{N} \tag{5}$$

where  $tp_c$  is the number of  $tp$  for class  $c$ ,  $C$  is the number of classes and  $N$  is the total number of instances in the dataset.

Macro average is the simple mean of scores of all classes and it is calculated as follows:

$$\beta_{macro} = \frac{1}{q} \sum_{\lambda=1}^q B(tp_{\lambda}, fp_{\lambda}, tp_{\lambda}, fn_{\lambda}) \tag{6}$$

Micro average or the weighted average is the sum of the marks of all classes after multiplying their class portions and it is calculated as follows:

$$\beta_{micro} = B\left(\sum_{\lambda=1}^q tp_{\lambda}, \sum_{\lambda=1}^q fp_{\lambda}, \sum_{\lambda=1}^q tn_{\lambda}, \sum_{\lambda=1}^q fn_{\lambda}\right) \tag{7}$$

The  $L = \lambda_j, (j = 1 \dots q)$  is the set of labels, where  $B(tp, tn, fp, fn)$  is evaluated centered on the number of  $tp, tn, fp,$  and  $fn$ , respectively.

Let  $tp_{\lambda}, fp_{\lambda}, tn_{\lambda}, \wedge fn_{\lambda}$  denote the number of  $tp, fp, tn,$  and  $fn$  after a binary estimate for a label  $\lambda$ .

Figure 6 shows the confusion matrix of the classes predicted by the Cloud-MobiNet model. The confusion matrix of the Cloud-MobiNet predictions in Figure 6 clearly shows our model’s confusion in classifying the clouds. For instance, confusing the cloud type of Ns and St is understandable for the reasons that they are very identical in shape, structure, transparency, and arrangement. The only difference that meteorological experts can sometimes use to distinguish them is their height, since the cloud type of Ns is a medium cloud, while the cloud type of St is a low cloud. The non-zero off-diagonal elements (0.1, 0.1, and 0.2) in the confusion matrix represent the percentage of the few misclassifications of the model.

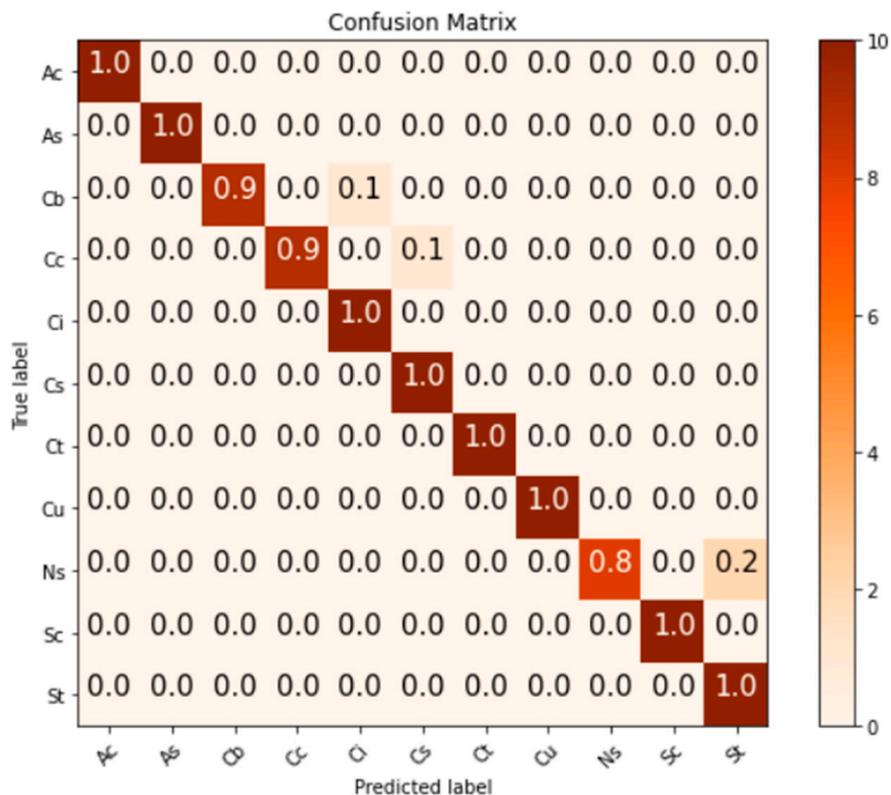


Figure 6. Confusion matrix.

### 3.1. Cloud-MobiNet Model’s Predictions and Interpretations

A prediction is an array of  $N$  numbers in which the  $N$  represents the number of classes (labels or categories) in the dataset. Each element represents the confidence that the image corresponds to in each of the image’s different classes (labels or categories). Figure 7 shows the predicted cloud, cloud class, and array.

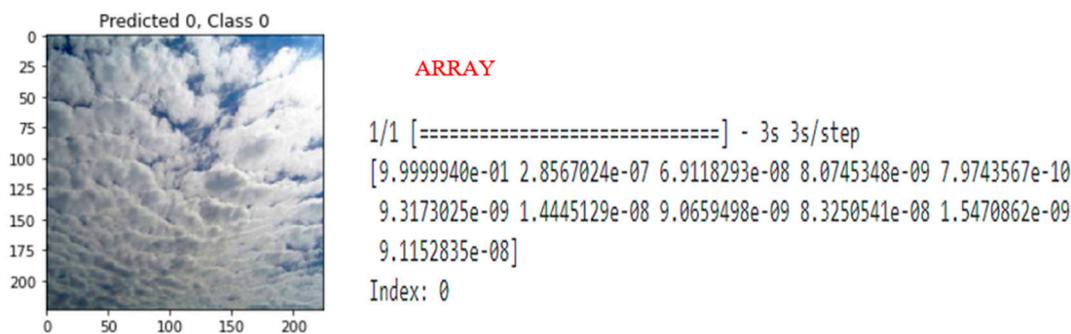


Figure 7. Sample predicted cloud, class, and array.

Based on 11 classes (0~10) from the reference sample picture in Figure 7, the model generated a prediction on index 0, since it had the highest confidence level of 99.99%, that the image class was 0. The model’s percentage confidence for each class on the image sample provided in Figure 7 is shown in Table 3.

Table 3. Outline of the model’s percentage confidence for the predicted Figure 8 sample image.

Label/Classes	Number (%)	Soil Name	Position
0	99.99%	Altocumulus	1st
1	0.000000286%	Altostratus	2nd
2	0.0000000691%	Cumulonimbus	5th
3	0.00000000807%	Cirrocumulus	9th
4	0.00000000797%	Cirrus	11th
5	0.00000000932%	Cirrostratus	7th
6	0.0000000144%	Contrails	6th
7	0.0000000907%	Cumulus	8th
8	0.0000000833%	Nimbostratus	4th
9	0.0000000155%	Stratocumulus	10th
10	0.0000000912%	Stratus	3rd

We randomly reserved 110 samples of the initial 2543 CCSN cloud dataset for testing and then performed the augmentation processes on the remaining 2433 cloud samples to have a unique test dataset not used for modeling. Out of 110 cloud images given to the model to predict their various classes, the model successfully predicted 106 correctly. Figure 8 shows some of the images of the clouds predicted by the Cloud-MobiNet model.

Each image has the name of the cloud (class) predicted by the model, the percentage confidence that the model attributed to each of the classes, and the actual cloud’s name. The graph beside each image shows how the model predicted the cloud images. A correctly predicted cloud image has a green bar, green predicted cloud name, and green actual cloud name. The red bar, red predicted cloud name, and red actual cloud name are those that the model failed to predict correctly. Images with more than one bar are an indication of the percentage confidence that the model gave to the image according to the features shared by the other classes.

### 3.2. Implementation of Cloud-MobiNet Model on Smartphone

Several types of Android building application software can be used to build and publish an Android app, such as the Cloud Prediction app. A software engineer’s or developer’s decision to choose a specific Android software-development application is determined by their expertise and familiarity with the program they select, as well as how comfortable they are with it. Because the model is not software-specific, it is capable of running on any preferred platform.

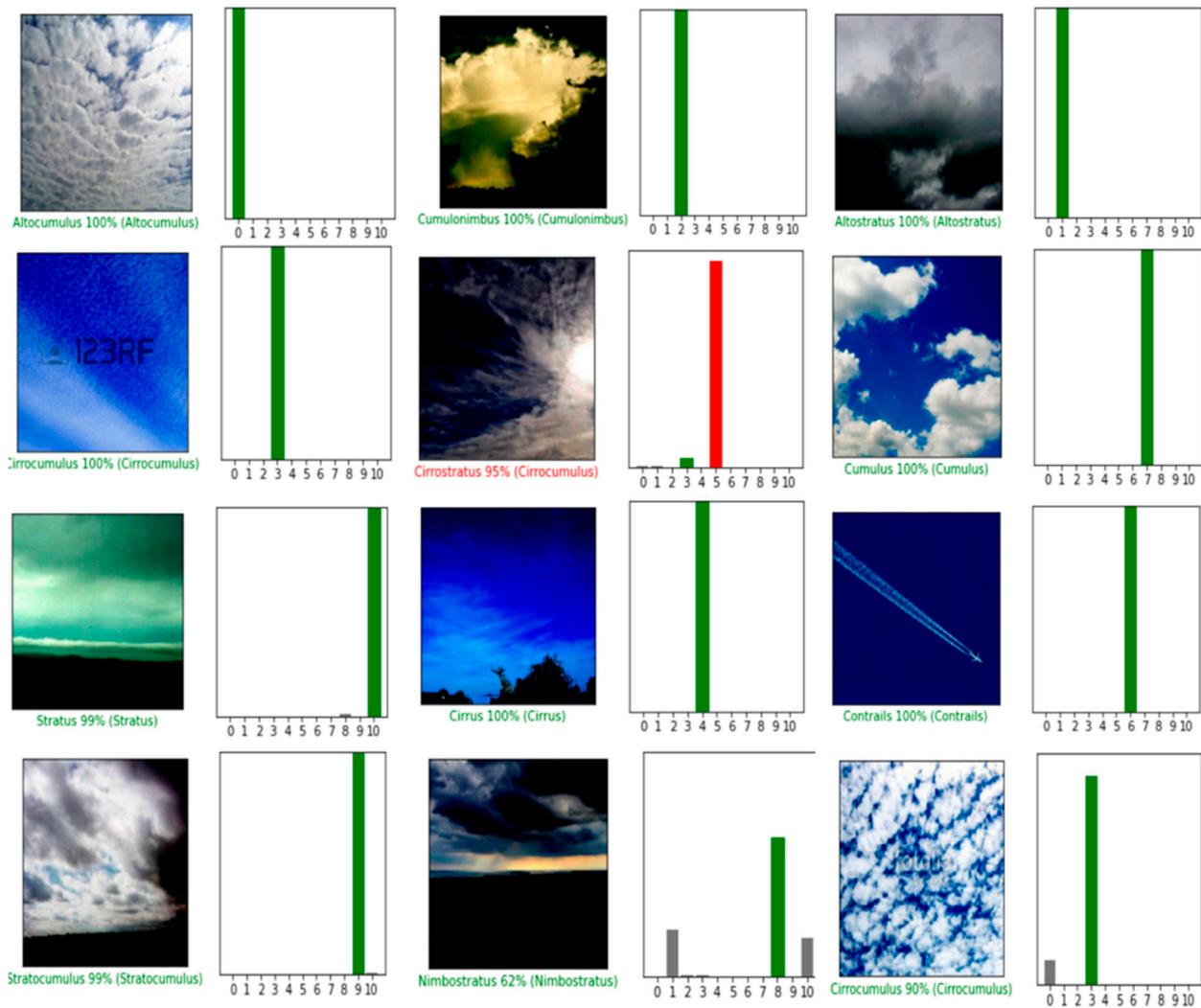


Figure 8. Samples of predicted cloud images.

The most basic interface-design requirement is to provide users with the option of using the camera of their smartphone to capture real-time images of clouds or to select a cloud image from their storage space. Figures 9 and 10 show the smartphone’s cloud-prediction-app progressions and the basic flowchart for implementing a cloud-prediction app on a smartphone using the Cloud-MobiNet model, respectively.



Figure 9. Smartphone’s cloud prediction app progressions.

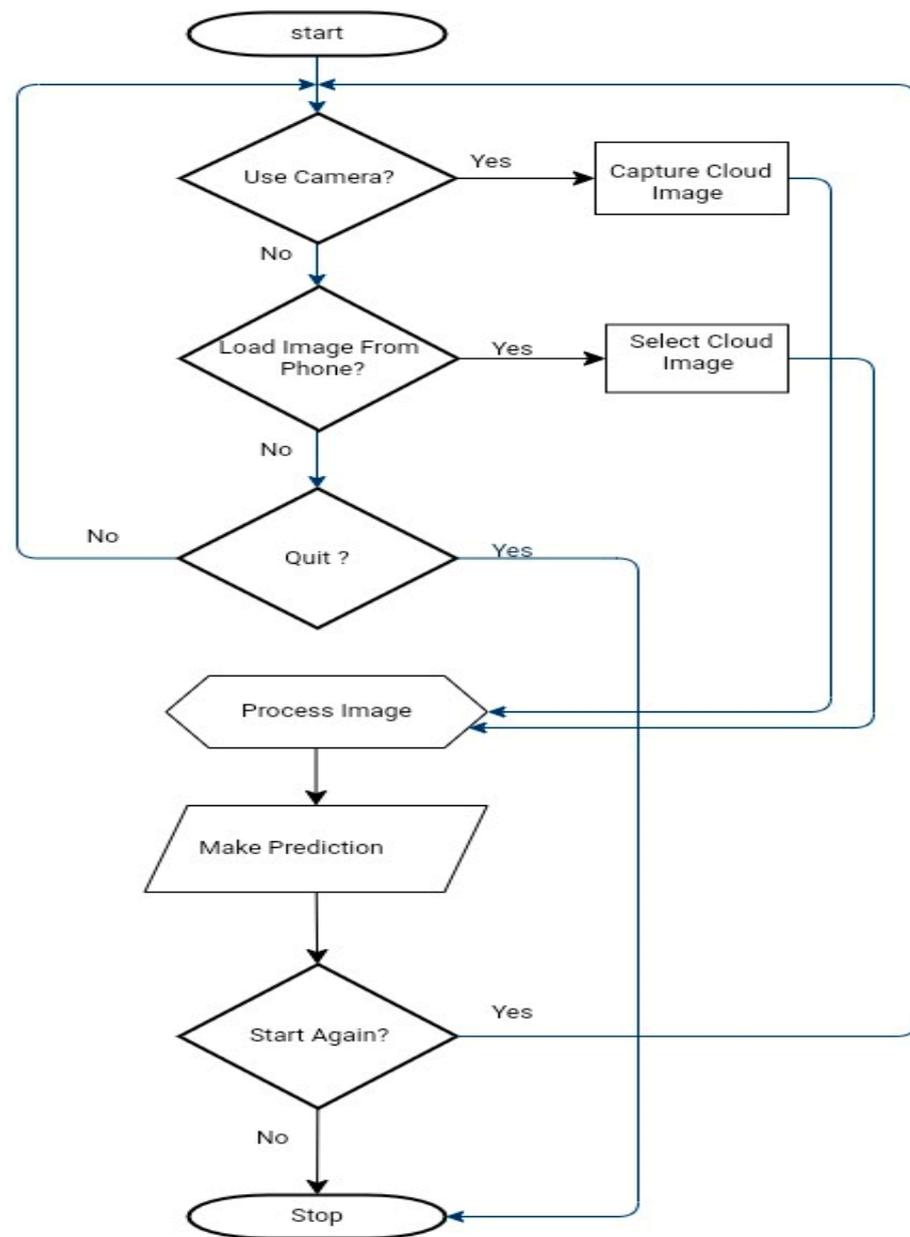


Figure 10. Flowchart of the smartphone implementation process of the Cloud-MobiNet model.

#### 4. Discussion

The classification report reveals the Cloud-MobiNet model's performance in identifying clouds in the CCSN dataset, as shown in Table 2. The cloud type of *St* has an excellent recall of 100%, and a good accuracy of 83%. The accuracy of the remaining cloud classifications ranges from 91% to 100%. This is a strong indicator of how well our Cloud-MobiNet model fits the problem, as it exhibits strong generalization, which eliminates overfitting and under-fitting during training, robustness, and the ability to accurately classify clouds. The performance of our Cloud-MobiNet model exceeds that of the much-touted CloudNet [21] and ultramodern techniques [5,10,24], with average scores of 88%, 87%, 81%, and 95%, respectively. The average score of 96% recorded by Cloud-MobiNet is a clear demonstration of how effectively the model can generalize the classification of the 11 classes of clouds. Table 4 is the comparison of our model's performance with the latest technological approaches in the literature.

**Table 4.** Comparison of this study with the latest technological approaches in the literature.

Article	Dataset	Year	Model/Method	Accuracy (%)
Zhang et al. (2018a)	CCSN	2018	CloudNet	88.0
Li et al. (2022)	ASGC,	2022	Transformer	94.2
	CCSN			92.7
	GCD			93.5
Zhu et al. (2022)	MGCD	2022	Combined convolutional network	90.0
	NRELCD			95.6
Fabel et al. (2022)	All sky images (Owned)	2022	Self-supervised learning	95.2
<b>Ours</b>	<b>CCSN</b>	<b>2023</b>	<b>Cloud-MobiNet</b>	<b>97.45</b>

The higher accuracy obtained by the model is a result of the robust dataset we generated and the smart model we constructed. In addition to its excellent performance, the Cloud-MobiNet model is also efficient, as only 130 epochs were run with a batch size of 22. This saves a significant amount of computational cost and resources, such as memory, time, and energy, compared to CloudNet, which runs 20,000 epochs with a batch size of eight.

## 5. Conclusions

Several deep-learning models have now emerged as a result of continued research in the area of artificial intelligence and neural networks. Based on MobileNet's convolutional neural network model, we propose a highly efficient deep-learning model called Cloud-MobiNet for the classification of ground-based cloud observation images. Cloud-MobiNet promises to be a significant model in the short term, since automated ground-based cloud classification is anticipated to be a preferred means of cloud observation, not only in meteorological analysis and forecasting but also in the aeronautical and aviation industries.

In addition to its compact and portable properties, the Cloud-MobiNet model's strength is that it can be used on mobile phones for real-time cloud classification. Contrary to the claim made by [21] that the CCSN dataset is complicated, Cloud-MobiNet has proven its supremacy in classifying the CCSN cloud dataset with a training and validation accuracy of 97.45 percent and an average testing accuracy of 96 percent, which is an improvement on CloudNet. Without ignoring its misclassifications of *Ns* and *St*, as well as *Cc* and *Cs*, the model is apt, since these pairs of clouds are almost not mutually exclusive in terms of texture, structure, and shape. We believe that with continued training, the model may attain an optimal accuracy of around 99 percent.

Even though the model was run on a regular laptop computer, the study was unusual in that it was not conducted in a controlled setting, nor was it based on camera sensors, nor was it altered or boosted by lighting conditions, as in previous studies undertaken by many experts. All of the resources needed, including energy (power), memory, speed, and disc space, are within the capabilities of any normal smartphone. Because the model is small enough, it was run on a laptop with standard hardware rather than a sophisticated server, making it easier to reproduce on any standard smartphone.

To implement this, the model must be compiled and executed as a mobile app. The cloud image must be accepted as input by the app's UI. The user may either load the image from a memory source, or utilize the phone's camera to capture the cloud image in real-time using the mobile app. The app with the model at the backend will process the image and then predict the type of cloud that is captured or loaded from the memory in real-time. We anticipate that after the model has been trained and tested on a variety of cloud images in various lighting conditions and has delivered extremely accurate results, the type of camera utilized by the cell phone will have little or no influence on the predictions.

Although the Cloud-MobiNet model predicts only the type of cloud, the identification of cloud types is the most difficult process in cloud observation for meteorologists. In this process, a wrongful determination of the type of cloud can have a strong influence on

weather forecasting. However, meteorologists know the average height of each of the 11 types of clouds and, for this reason, obtaining the correct cloud type helps the observers to determine their height. In weather forecasting, the cloud type is critically examined more than the height. It is also easy for meteorologists to determine the cloud cover in the sky at any time. Future work will concentrate on how feasible the model can be used to determine both the height and cloud amount.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/atmos14020280/s1>, Supplementary Document: SD1\_Cloud-MobiNet\_Codes.

**Author Contributions:** E.K.G.: Conceptualization, writing—original draft; methodology, visualization, writing—review and editing, software, formal analysis, data curation, resources. P.S.: funding acquisition, writing—review and editing; supervision; investigation; validation; project administration. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Vellore Institute of Technology (VIT), grant number “VIT SEED GRANT” and the APC is funded by VIT University.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The dataset is available at <https://github.com/upuil/CCSN-Database>.

**Acknowledgments:** We express our deepest gratitude to Vellore Institute of Technology (VIT) University for their support, and Zhang, J., for making the CCSN dataset available to us.

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

## References

1. Duda, P.; Minnis, P.; Khlopenkov, K.; Chee, T.L.; Boeke, R. Estimation of 2006 Northern Hemisphere contrail coverage using MODIS data. *Geophys. Res. Lett.* **2013**, *40*, 612–617. [CrossRef]
2. Kim, I.; Martins, R.J.; Jang, J.; Badloe, T.; Khadir, S.; Jung, H.Y.; Kim, H.; Kim, J.; Genevet, P.; Rho, J. Nanophotonics for light detection and ranging technology. *Nat. Nanotechnol.* **2021**, *16*, 508–524. [CrossRef] [PubMed]
3. Calbo, J.; Sabburg, J. Feature extraction from whole-sky ground-based images for cloud-type recognition. *J. Atmos. Ocean. Technol.* **2008**, *25*, 3–14. [CrossRef]
4. Zhuo, W.; Cao, Z.; Xiao, Y. Cloud classification of ground-based images using texture, and structure features. *J. Atmos. Ocean. Technol.* **2014**, *31*, 79–92. [CrossRef]
5. Xiao, Y.; Cao, Z.; Zhuo, W.; Ye, L.; Zhu, L. mCLOUD: A multi-view visual feature extraction mechanism for ground-based cloud image categorization. *J. Atmos. Ocean. Technol.* **2016**, *33*, 789–801. [CrossRef]
6. Kazantzidis, A.; Tzoumanikas, P.; Bais, A.F.; Fotopoulos, S.; Economou, G. Cloud detection and classification with the use of whole-sky ground-based images. *Atmos. Res.* **2012**, *113*, 80–88. [CrossRef]
7. He, T.; Zhang, Z.; Zhang, H.; Zhang, Z.; Xie, J.; Li, M. Bag of tricks for image classification with convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 558–567.
8. Labati, R.D.; Muñoz, E.; Piuri, V.; Sassi, R.; Scotti, F. Deep-ECG: Convolutional neural networks for ECG biometric recognition. *Pattern Recognit. Lett.* **2019**, *126*, 78–85. [CrossRef]
9. Shi, C.; Wang, C.; Wang, Y.; Xiao, B. Deep convolutional activations-based features for ground-based cloud classification. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 816–820. [CrossRef]
10. Ye, L.; Cao, Z.; Xiao, Y. Deep cloud: Ground-based cloud image categorization using deep convolutional features. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5729–5740. [CrossRef]
11. Liu, S.; Zhang, Z. Learning group patterns for ground-based cloud classification in wireless sensor networks. *Eurasip J. Wirel. Commun. Netw.* **2016**, *2016*, 69. [CrossRef]
12. Liu, S.; Li, M.; Zhang, Z.; Cao, X.; Durrani, T.S. Ground-Based Cloud Classification Using Task-Based Graph Convolutional Network. *Geophys. Res. Lett.* **2020**, *47*, e2020GL087338. [CrossRef]
13. Taigman, Y.; Yang, M.; Ranzato, M.A.; Wolf, L. Deepface: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1701–1708.
14. Schroff, F.; Kalenichenko, D.; Philbin, J. Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 815–823.

15. Lee, J.; Weger, R.C.; Sengupta, S.K.; Welch, R.M. A neural network approach to cloud classification. *IEEE Trans. Geosci. Remote Sens.* **1990**, *28*, 846–855. [[CrossRef](#)]
16. Bin, T.; Mukhtiar, S.; Sadjadi, A.; Hear, T.H.V.; Reinke, D.L. A study of cloud classification with neural networks using spectral and textural features. *IEEE Trans. Neural Netw.* **1999**, *10*, 138.
17. Fabel, Y.; Nouri, B.; Wilbert, S.; Blum, N.; Triebel, R.; Hasenbalg, M.; Kuhn, P.; Zarzalejo, L.F.; Pitz-Paal, R. Applying self-supervised learning for semantic cloud segmentation of all-sky images. *Atmos. Meas. Tech.* **2022**, *15*, 797–809. [[CrossRef](#)]
18. Li, X.; Qiu, B.; Cao, G.; Wu, C.; Zhang, L. A Novel Method for Ground-Based Cloud Image Classification Using Transformer. *Remote Sens.* **2022**, *14*, 3978. [[CrossRef](#)]
19. Toğaçar, M.; Ergen, B. Classification of cloud images by using super-resolution, semantic segmentation approaches and binary sailfish optimization method with deep learning model. *Comput. Electron. Agric.* **2022**, *193*, 106724. [[CrossRef](#)]
20. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNet: Efficient Convolutional Neural Networks for Mobile Applications. *arXiv* **2017**, arXiv:1704.04861.
21. Zhang, J.L.; Liu, P.; Zhang, F.; Song, Q.Q. CloudNet: Ground-based cloud classification with a deep convolutional neural network. *Geophys. Res. Lett.* **2018**, *45*, 8665–8672. [[CrossRef](#)]
22. Szegedy, C.; Vanhoucke, V.; Ioe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
23. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv* **2016**, arXiv:1603.04467.
24. Zhu, W.; Chen, T.; Hou, B.; Bian, C.; Yu, A.; Chen, L.; Tang, M.; Zhu, Y. Classification of Ground-Based Cloud Images by Improved Combined Convolutional Network. *Appl. Sci.* **2022**, *12*, 1570. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.