Method for Classification of Snowflakes Based on Images by a Multi-Angle Snowflake Camera Using Convolutional Neural Networks

A. Hicks and B. M. Notaroš†

Department of Electrical and Computer Engineering

Colorado State University, Fort Collins, CO, USA

Submitted to **Journal of Atmospheric and Oceanic Technology**

April 7, 2019

Revised Manuscript, 21 July 2019

†Corresponding Author:
Branislav M. Notaroš
Colorado State University
Department of Electrical and Computer Engineering
1373 Campus Delivery
Fort Collins, CO 80523, USA
Phone: (970) 491-3537, Fax: (970) 491-2249
Web: www.engr.colostate.edu/~notaros
E-mail: notaros@colostate.edu

Early Online Release: This preliminary version has been accepted for publication in *Journal of Atmospheric and Oceanic Technology*, may be fully cited, and has been assigned DOI 10.1175/JTECH-D-19-0055.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

ABSTRACT

Taking advantage of the recent developments in machine learning, we propose an approach to automatic winter hydrometeor classification based on utilization of convolutional neural networks (CNNs). We describe the development, implementation, and evaluation of a method and tool for classification of snowflakes based on geometric characteristics and riming degree, respectively, obtained using CNNs from high-resolution images by a Multi-Angle Snowflake Camera (MASC). These networks are optimal for image classification of winter precipitation particles due to their high accuracy, computational efficiency, automatic feature extraction, and application versatility. They require little initial preparation, enable the use of smaller training sets through transfer learning techniques, come with large supporting communities and a wealth of resources available, and can be applied and operated by non-experts. We illustrate both the ease of implementation and the usefulness of operation the CNN architecture offers as a tool for researchers and practitioners utilizing in-situ optical observational devices. A training data set containing 1,450 MASC images is developed primarily from two storm events in December 2014 and February 2015 in Greeley, Colorado, by visual inspection of recognizable snowflake geometries. Defined geometric classes are aggregate, columnar crystal, planar crystal, small particle, and graupel. The CNN trained on this data set achieves a mean accuracy of 93.4% and displays excellent generalization (ability to classify new data). In addition, a separate training data set is developed sorting snowflakes into three classes showcasing distinct degrees of riming. The CNN riming degree estimator yields promising initial results but would benefit from larger training sets.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

1. Introduction

The advent of dual polarimetric radar for weather observation and research has increased our capabilities to access and log vital data points within a given weather event. With excellent temporal and spatial resolution, researchers can accurately characterize shapes of the hydrometeors that compose a storm (Straka et al. 2000). Polarimetric radars provide the horizontal reflectivity, Z_{h}, differential reflectivity, Z_{dr}, and correlation coefficient, ρ_{hv}, of a field the radar is directed toward. This information gives insight to shape and type of hydrometeors within a storm and is calculated based on models developed by past observations. Scattering models fall short when left to spheroid approximations for frozen hydrometeors (Tyynelä et al. 2011). This is especially true at higher frequencies or for larger particles (Kim 2006), and increases the need for accurate accounting of the varying microphysical characteristics of snow within a storm. Atmospheric scientists have drawn strong correlation between the environmental conditions present and the shape snow takes as it forms within a storm (Libbrecht 2017). Conditions within a storm are not homogenous, resulting in a wide variety of shapes that continue to change on the hydrometeor's path to the ground. While polarimetric radar provides excellent coverage, utilizing ground based (in-situ) devices in tandem with radar has proven more effective in understanding storm composition (Zhang et al. 2011), especially in studying riming degree. Riming (the collection of supercooled water droplets onto an ice crystals surface) is one of the physical metrics that can indicate valuable information about the internal characteristics of a storm and is of significant interest to the atmospheric science community (e.g., Kennedy et al. 2018). In-situ devices are often deployed on the ground in the path of a storm and allow for detailed sampling utilizing high resolution imaging techniques to capture individual hydrometeors while monitoring local environmental conditions (e.g.,

Accepted for publication in Journal of Atmospheric and Oceanic Technology. DOI 10.1175/JTECH-D-19-0055.1.

temperature, humidity, and wind speed). Some examples of in-situ image capturing devices include the Two-Dimensional Video Disdrometer (2DVD; Schönhuber et al. 2008) or the Precipitation Instrument Package (PIP; a more advanced version of the Snow Video Imager described in Newman et al. 2009). Another such in-situ device developed specifically to sample snowflakes in free fall is the Multi-Angle Snowflake Camera (MASC). The MASC can capture high resolution images of individual snowflakes, which provides researchers an avenue to study the microphysical characteristics and make statistical predictions concerning a storm's composition. A deployed MASC is capable of capturing thousands of images an hour, and with typical storms lasting several hours, a fast, accurate and automatic method to organize and process image data based on shared characteristics is crucial to increased understanding. There exist several classification algorithms known to atmospheric research communities (e.g., Chandrasekar et al. 2013; Besic et al. 2016) but they are limited to large swaths of a storm and not local sampled images. Developing a classifier that functions automatically on a per snowflake basis would be a critical first step in introducing intelligent post-storm processing capabilities to in-situ devices, reducing data processing time for research, and allowing devices to remain in the field for longer periods.

Machine learning algorithms have made huge strides in the past decade (Minar and Naher 2018), especially with classification tasks, and have a massive community invested in improving and expanding existing algorithms. Previous attempts have been made to apply machine learning to snowflake classification with varying degrees of success. Early backpropagation neural networks (BPNNs) were used by Feind (2006) and multinomial logistic regression (MLR) by Praz et al. (2017). Feind (2006) achieved their best results using BPNNs with 85% accuracy in classifying eight categories of hydrometeors (drops, snow, hail, columns,

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

needles, plates, dendrites, and holes). Data was extracted using a PMS 2D-C probe that creates detailed particle profiles and is mounted on a T-28 aircraft which flies directly through storms. The images from the PMS 2D-C do not allow for microphysical characteristics of the particles to be considered, as the data is black and white profiles. Praz et al. (2017) set the standard for expert research into hydrometeor classification with a separate algorithm for feature extraction and optimization, and achieved a classification accuracy of 95%. Due to the high-resolution images used as their sample data, they can consider particle microphysical characteristics, as well as make estimations regarding degree of riming present. On the other hand, their algorithm requires external feature extraction and optimization, a process that may limit input complexity and number of classes due to the computational inefficiencies within the architecture. The feature extraction process would also need to be repeated with any hardware change as they likely incorporate hardware bias (imperfections unique to the device) into a data set, thus reducing the generality and versatility of the method.

This paper takes advantage of the most recent developments in machine learning and proposes an approach to automatic winter hydrometeor classification based on utilization of convolutional neural networks (CNNs). It presents the development, implementation, and evaluation of a CNN-based method and tool for classification of snowflakes using high-resolution images by a Multi-Angle Snowflake Camera. Snowflake classifications based on geometric characteristics and riming degree, respectively, are described and tested. Convolutional neural networks by their nature and properties are an excellent candidate for an algorithm and tool for classification of hydrometeors (and particularly winter precipitation) based on high-resolution images of particles. They were developed with image processing in mind, which makes them computationally more efficient for image-based classification when

Accepted for publication in Journal of Atmospheric and Oceanic Technology. DOI 10.1175/JTECH-D-19-0055.1.

compared to other multilayer backpropagation neural networks. CNNs act as a sort of "black box", which automatically extract features during training, thus simplifying any system they are integrated into, with interest in data specific metrics available through further processing. CNNs can store these features, which increases their versatility as they are capable of transferring learning from one data set to another and are not limited to specific parameters inherent either to the data set (e.g., resolution, color, or size) or capturing method (e.g., hardware imperfections reflected in data). Therefore, a classifier properly trained with a CNN can be utilized by a variety of image-capturing in-situ devices. Research into deep learning has extended their ability to process complex data without major changes to the algorithm. Finally, CNNs are well understood algorithms that are extremely popular for image processing with a wealth of resources available, thus reducing a reliance on expert help in implementation (Mathworks 2018a, 2018b).

In this paper, the steps needed to develop a hydrometeor classifier using CNNs are presented in detail and advantages to in-situ research are highlighted. The intention of this work is to illustrate the ease of implementation the CNN architecture offers as a tool for researchers and practitioners utilizing in-situ measurement devices. The CNN method described here requires less training data [e.g., 1,450 training samples, compared to 2,000 used by Fiend (2006) or 3,000 used by Praz et al. (2017)], as well as less image preprocessing, while attaining a geometric classification accuracy of 93.4%, for instance, which is comparable to other classifiers. Moreover, the new method provides extra flexibility for expanded functionality (e.g., additional classes, different hydrometeor types, etc.) that can readily and non-expertly be achieved as the backend of a deployed measurement device and frontend to further data processing and analyses. Data preprocessing is reduced to cropping images to remove instances

Accepted for publication in Journal of Atmospheric and Oceanic Technology. DOI 10.1175/JTECH-D-19-0055.1.

where multiple snowflakes are present and a brightness thresholding filter to remove images which are too dim or blurry. Both steps are handled by a simple script which can then feed the processed images directly to each classifier. The classifier then organizes the data based on image features and places data in folders labeled for each class. This process is fast and computationally efficient; for example, the tool can be deployed in computers typically running in-situ measurement devices. The classifier in this work is developed using MATLAB™ 2018, but open-source toolboxes are available if additional flexibility is required.

2. Snowflake Data Collection and Preparation

This section describes how and under what conditions data was collected. Steps required in preprocessing of image data are covered. The classification criteria for geometric shapes and riming degree estimation are discussed.

2.1 Data Collection Site

The images that compose the training set were taken primarily from two winter weather events using a modified MASC system located at a surface instrumentation field site that was established as part of MASCRAD (MASC + RADar) project (Notaroš et al 2016; Bringi et al. 2017; Kennedy et al. 2018). The MASCRAD field site, Figure 1**,** is located at the Easton Valley View Airport, in La Salle, outside of Greeley, Colorado. This site includes a modified MASC system, a 2DVD, a PLUVIO all-weather precipitation gauge, and a VAISALA weather station, amongst other advanced in-situ measurement instruments. These devices are situated within a double fence intercomparison reference (DFIR) and operate under the umbrella of the CSU-CHILL Radar, a state-of-the-art polarimetric weather radar located 12.92 km away.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

The image data was collected during two events in the 2014-2015 deployment season of the MASCRAD project. The first took place from December 23rd to 31st in 2014, with the highest density of particles falling during the early morning hours of December 26th. To include enough graupel images, a second weather event was required (Bang et al. 2016). This event took place on February 21st – 22nd in 2015.

2.2 Data Collection Device

The Multi-Angle Snowflake Camera, or MASC, is the centerpiece of the MASCRAD project. While the finer details of the system are provided in (Garrett et al. 2012), for the purposes of this study a summary and brief description follows. The original MASC system is an instrument used to capture high-resolution images and fall speeds of hydrometeors in freefall from three different coplanar perspectives. These cameras are 5 Megapixel (MP) Unibrain Fire-I 980b digital cameras with identical 12.5-mm Fujinon lenses. The cameras are spaced on a horizontal ring with 36° separation between adjacent cameras, with camera-to-common focal center distances of 10 cm. The system used at Colorado State University, Figure 2, has been modified to include two additional cameras at an elevated angle of 55° above the horizon. These are 1.2 MP Unibrain Fire-I 785b cameras with 12.5-mm lenses, included to improve the 3D virtual reconstruction using the visual hull method (Kleinkort et al. 2017). The system has a horizontal resolution of 35 μm for the three horizontal cameras and a vertical resolution of 40 μm at 1-m/s fall speed. As hydrometeors fall through the horizontal ring, a near-IR emitter-detector pair sensor array (located on the top rim of the capture volume within the ring) simultaneously triggers the cameras and a flash (LEDs). The cameras have a maximum triggering rate of 2 Hz, a hardware limitation within the cameras, not set by the emitter-sensor

Accepted for publication in Journal of Atmospheric and Oceanic Technology. DOI 10.1175/JTECH-D-19-0055.1.

pair array. Finally, measurement of the time between upper and lower near-IR emitter-sensor triggers by a particle is used to calculate the particle fall speed (which is not the topic of this work).

2.3 Image Preprocessing

As a CNN automatically extracts features (numerical descriptors common between classes) during training that are to be used in classification, the image data base requires minimal preprocessing. A simple brightness thresholding is utilized to remove the majority of blank, dim or blurry images. Generally, a MASC is deployed during storming conditions, those with heavy wind or flurried snow, which can (in spite of the DFIR) cause the MASC to trigger without a snowflake in the focal area, making the need for thresholding crucial to processing. After thresholding to remove poor quality images, the creation of a training set requires a one-to-one correlation between an image of a snowflake and the class the image is being assigned to, i.e., there can only be one snowflake per image. The task of separating images was automatically performed using a cropping script developed to find the brightest point of an image, locate the surrounding edges through their calculated standard deviations, and remove the snowflake to be saved in another location. The script then performs this action again until all independent bright spots have been cropped in a given image. This procedure is not perfect (<10% of snowflakes need repairs) but is dramatically more efficient than cropping images by hand.

3. Hydrometeor Classification Scheme and Training Sets

With machine learning algorithms, there are two categories of how learning is conducted: supervised or unsupervised. CNNs are supervised learning algorithms, therefore they require the

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

development of a labeled data set, referred to as the training set. This training set allows a human operator to dictate how the network makes decisions, by providing desired data for the network to make comparisons against. This contrasts unsupervised learning, where a network runs until it converges on a pattern or an end condition is reached. As expected, supervised learning is more efficient but does require additional setup in the development of the learning set and may be subject the supervisor's bias in selecting representative data.

The training set is developed by human inspection following a predetermined classification scheme. The idiom is that every snowflake is unique, therefore it is no surprise that there are a variety of attempts to classify them (Korolev and Sussman 2000, Grazioli et al. 2014), with little commonality between schemes. The scheme used in this work was adopted from Praz et al. (2017), who developed a snowflake classifier using Multinomial Logistic Regression (MLR) and will be summarized in this section.

The scheme utilizes the nine categories of snowflakes detailed in (Magono and Lee 1966), with some changes due to data availability and simplification. They introduce the category of Aggregate particles, which are single snowflakes that are the result of the in-air collision of two or more particles and Small Particles, snowflakes whose feature characteristics are too small to categorize. They also combine the category of needle and column type snowflakes, as they share similar characteristics. Their result was 10 individual categories that include Aggregate (AG), Small Particle (SP), Columnar Crystal (CC; the resulting combination of needle and column particulates), Planar Crystal (PC), a Combination of Columnar Crystals, a Combination of Planar Crystals, Combination of Columnar and Planar Crystals, Graupel (GR), Irregular Snow Crystal, and Germ of Snow. Due to limited sample representation, the training set used in this paper includes only the most populated five classes, AG, CC, PC, SP, and GR,

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

shown in Figure 3, although more classes may be added to the classifier as sample data is accrued. A different training set is used for each classifier, and therefore their development considerations are unique to each set.

3.1 Geometric Training Set

At the heart of supervised learning is the comparison between the desired value (established in the training set) and a value calculated by the network, that is later refined through the learning process. This is an oversimplification, with more in-depth analysis provided in Section 4, but it is important to understand that the more unique the classes within a training set are from each other, the better the network will perform, given a limited training set (which seems obvious but cannot be overstated). The training set for geometric classification was developed by selecting individual images of the snowflakes that best represented their respective classes, with decision emphasis placed on discernable snowflake silhouettes. The result is a training set of ~1,450 samples.

3.2 Riming Degree Estimation Training Set

Riming degree estimations attempt to calculate the amount of cloud frozen droplets that accrue on a snowflake's surface as the snowflake falls through the atmosphere. The MASC system captures images with sufficiently detailed resolution that the degree of riming can be considered a feature of the image. There are two approaches to riming degree estimation considered in this paper. The first approach divides riming degree into 5 classes utilizing the classification scheme adapted by Praz et al. (2017). The second approach is a proposal that capitalizes on the unique nature of CNNs and warrants further exploration.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology*. DOI 10.1175/JTECH-D-19-0055.1.

Praz et al. (2017) classify riming degree based on the image criteria summarized in Table 1. These five degrees of riming, R_d, are discrete classes between [1,5] as developed by (Mosimann et al. 1994) and are then mapped by Praz et al. (2017) to a continuous index, R_c, between [0,1], using a sinusoidal function:

$$R_c = \frac{1}{2}\left(\sin\left(\frac{\pi}{4}(R_d - 3)\right) + 1\right). \quad (1)$$

The degree label decisions are based on educated opinion through the observation of a captured image. Questions arise when pondering the level of accuracy that a human observer can achieve in their estimation. For example, the difference between a riming degree estimation of 3.1 and 3.2 is often arbitrary and a matter of opinion or even capability. To address this concern, it is proposed to utilize the posterior distribution (a result of the classifier) that a CNN uses to make classification decisions as the deciding factor in riming degree estimation. The posterior distribution is a numerical value the classifier assigns to each snowflake describing the probability that it belongs to a given class or label. If labels are restricted to discrete values, the probability can be interpreted as the likelihood that a snowflake falls somewhere on that scale, making a continuous estimation.

Developing the training set for this approach, the classification scheme focuses on the three easily identified classes for riming estimation, then allows the network to assign probability estimations (clarified at the end of Section 4.2) for how closely a snowflake resembles those classes, labeled $R_{l,c}$ and characterized in Table 1. The estimate training classes are class 1, where snowflakes are the least rimed (no riming present), class 3, where a snowflake is rimed but geometric shape is preserved, and class 5, where the snowflake is fully rimed (graupel). This simplifies the classification process and increases the consistency in decision making when applying labels while developing the training set. The result after classification is snowflakes that

Accepted for publication in Journal of Atmospheric and Oceanic Technology. DOI 10.1175/JTECH-D-19-0055.1.

do not fall directly within these three classes are weighted somewhere between. For example, if a snowflake's riming degree estimation is $R_{l,c1} = 32\%$ $R_{l,c3} = 58\%$ $R_{l,c5} = 10\%$, the estimation on Praz et al.'s (2017) scale is akin to ~2.78. The numerical values are used as an example, with a simple linear mapping. As more images are processed, a more accurate mapping can be developed if desired. This provides a continuous riming degree estimation and the opportunity for greater accuracy in estimation. Attention must be paid to the number of geometric classes represented in each riming category, as it is important to have uniform representation. This restriction limits the size of training set available, but the approach looks promising and will warrant further testing as processed data becomes available.

4. Convolutional Neural Networks Method and Code

A brief discussion of the network architecture is presented in this section. An overview of key concepts pertaining to CNNs and deep learning is provided. The software implementation and input parameters outlined before results of the network training are discussed.

4.1 Neural Network Architecture

A human brain can process information very quickly, namely, it has an ability to rapidly take incoming information, assign meaning, and make decisions. This is accomplished through a complex interconnected system of neurons that process information in parallel. It is no surprise then that considerable effort has been made to mimic the complex way the brain processes information with machine learning algorithms.

With artificial neural networks, more commonly known as "neural networks," the basic architecture of the brain is recreated in a logical algorithm. The basis for the network architecture

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

used in this paper is called the "Multi-Layer Perceptron (MLP)," which is a feed forward neural network with backpropagation, whose general structure is represented in Figure 4a, and as an algorithm in Figure 4b. In general, there is an input signal which is passed to the hidden layers for processing, with final decision making the result of the output layer. The algorithm is "feed forward" in reference to the direction of information and the "backpropagation" describes the learning process, which is detailed below.

Referring to Figure 4b, the forward pass of the learning algorithm applies the input value $x_{1...N}$ of training sample *p* with a weighted bias of $\mathrm{w}_{kj}^{(1)}(p)$. This bias serves to both connect the input *j* to neuron *k* and to refine the error estimation of the classifier, E_p, and are generally random values between a range loosely defined by the number of inputs (Azimi 2018). The potential of neuron *k* is represented by

$$u_k^{(1)}(p) = \sum_{j=1}^{N+1} \mathrm{w}_{kj}^{(1)}(p) x_j(p), \quad \forall k \epsilon [1, K]. \tag{2}$$

The nonlinear activation function is, commonly, the logistic function $f(\cdot)$ (Murphy 2012) or, more recently (and in the case of this paper), the Rectified Linear Unit (ReLU),

$$f(u) = \frac{1}{1+\exp(-u)} \quad \text{or} \quad ReLU(u) = max(0, \mathrm{u}). \tag{3}$$

This activation function determines the output of the current neuron, which, in turn, becomes the input of the next layer within the network,

$$o_k^{(1)}(p) = f\left(u_k^{(1)}(p)\right), \quad \forall k \epsilon [1, K]. \tag{4}$$

This process is repeated until the final layer of the network is reached, labeled "Output" in Figure 4b. It is at this point that the sum squared error of the classifier can be calculated,

$$E_p = \frac{1}{2} \sum_{m=1}^{M} \left(d_m^{(output)}(p) - o_m^{output}(p)\right)^2. \tag{5}$$

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

Our goal is to minimize E_p, such that the desired output, $d_m^{(output)}(p)$, is as close to the calculated output, $o_m^{output}(p)$, as possible. It is unlikely that E_p is within acceptable parameters on the first pass of learning, therefore there needs to be some method to refine the weights to approach convergence. This requires calculation of the partial derivative of E_p with respect to the weight in the final layer, namely, $\mathrm{w}_{ml}^{(output)}$ in Figure 4b. The full derivation falls outside the scope of this paper but can be found in various forms in (Murphy 2012; Haykin 2009; Syozil et al. 1997). The final result for this example three-layer network is

$$\Delta\mathrm{w}_{ml}^{(output)}(p) = -\mu\nabla E_p(p) = \mu e_m^{output}(p) f'\left(u_m^{(output)}(p)\right) o_l^{H-1}(p),$$

$$\forall m \in [1, M] \text{ and } \forall l \in [1, L], \tag{6}$$

where

$$e_m^{output}(p) = d_m^{(output)}(p) - o_m^{output}(p), \quad \mu = learning\ factor. \tag{7}$$

Repeating this derivation for weight updates within the hidden layers shows that the calculated layer from the output is back-propagated through the network,

$$\Delta\mathrm{w}_{ml}^{(output)} = \mu e_m^{output}(p) f'\left(u_m^{(output)}(p)\right) o_l^{H-1}(p), \tag{8}$$

$$\Delta\mathrm{w}_{ml}^{(output)} = \mathrm{w}_{ml}^{(output)}(p+1) - \mathrm{w}_{ml}^{(output)}(p), \quad \forall m \in [1, M], \tag{9}$$

$$\Delta\mathrm{w}_{lk}^{(H-1)} = \mu f'\left(u_l^{(H-1)}(p)\right) o_k^1(p) \textstyle\sum_{m=1}^{M} \mathrm{w}_{ml}^{(output)}(p) e_m^{output}(p) f'\left(u_m^{(output)}(p)\right), \tag{10}$$

$$\Delta\mathrm{w}_{lk}^{(H-1)} = \mathrm{w}_{lk}^{(H-1)}(p+1) - \mathrm{w}_{lk}^{(H-1)}(p), \quad \forall l \in [1, L], \tag{11}$$

$$\Delta\mathrm{w}_{kj}^{(1)} = \mu f'\left(u_k^{(1)}(p)\right) x_j(p) \sum_{l=1}^{L} \mathrm{w}_{lk}^{(H-1)}(p)\, f'\left(u_l^{(H-1)}(p)\right)$$

$$* \left[\textstyle\sum_{m=1}^{M} \mathrm{w}_{ml}^{(output)}(p) e_m^{output}(p) f'\left(u_m^{(output)}(p)\right)\right], \tag{12}$$

$$\Delta\mathrm{w}_{kj}^{(1)} = \mathrm{w}_{kj}^{(1)}(p+1) - \mathrm{w}_{kj}^{(1)}(p), \quad \forall k \in [1, K]. \tag{13}$$

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

This back and forth is continued until the error stops decreasing or an acceptable value is reached.

4.2 Convolutional Neural Networks

With image classification, the input value of $x_{1 \ldots N}$ is a pixel of the image in the training set and this presents a major limitation to typical MLP architectures. Each input has an individual weight value per neuron in the network, so for even a moderately deep network (the number of hidden layers in the network refers to how deep a network is), the result is upwards of hundreds of thousands of weights that require refinement, or even millions for very high resolutions, making this computationally inefficient for modern practical purposes. CNNs solve many of the problems MLPs experience for image processing and have proven useful with other data types (e.g., Collobert and Weston 2008, O'Shea et al. 2016). A visual representation of the CNN architecture is shown in Figure 5.

The defining features of the CNN architecture are the inclusion of the convolution layer, a pooling layer, and the fully connected layer. The convolution operator introduces several advantages to the architecture, namely sparse interactions, parameter sharing and equivariance to translation (Goodfellow et al. 2016). As is depicted in Figure 5, the convolution layer creates a series of feature maps by scanning a weight matrix of size [i,j] over the surface of the input data. This reduces the number of parameters the network must consider and introduces weight sharing. Comparing to Figure 4b, the inputs are blocked or shared between units with shared weights. The result is a reduction in memory requirements and an improvement of statistical efficiency (Goodfellow et al. 2016). After the convolution is performed, the linear activations that result are then passed through a nonlinear activation function, like the ReLU which has largely replaced

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

other activation functions, as it improves efficiency without a reduction in accuracy (Goodfellow et al. 2016). The convolution layer can be repeated with different sized weight matrices to extract more and more abstract features. LeCun et al. (1998) introduced the pooling, or subsampling, layer to their model **LeNet5** to achieve shift invariance **(**Murphy 2012). This is accomplished by either averaging or producing a max over a small window of the convolution layer. This step is especially vital for image classification, as it allows the network to extract features without concerning exactly where the features are located (Goodfellow et al. 2016). An additional benefit of this property is that it makes transfer learning with pretrained networks possible (Torrey and Shavlik 2009), which significantly reduces the size of training set necessary for new classification schemes. **LeNet5** followed every convolution with a pooling layer, but this is unnecessary. It has been shown that the best results for complex data sets apply a few pooling layers after the first series of convolution layers and a final pooling layer after the next to last convolution layer (Romanuke 2017). The final few layers of the network will consist of fully connected layers which are akin to those in regular neural networks. The output of these layers after activation are passed to a *softmax* operator $\sigma(\cdot)$ where high-level decisions are made, and after several passes, a class label is applied. The learning process is summarized as

$$y_j = \sigma(\underline{o})_j = \frac{\exp(z_j)}{\sum_{k=1}^{K}\exp(z_k)}, \quad \forall j \in [1, K]\,, \tag{10}$$

where $y_j = p(C_j|\underline{x})$, a posterior distribution over the available classes, with C_j standing for the class and $\underline{x}$ for the input vector. In Section 3.2, a term “likelihood” is used to provide a label to the degree of riming present for a given snowflake, for added clarity, $y_1 = R_{l,c1}, y_3 = R_{l,c3}$, and $y_5 = R_{l,c5}$. If $\mathcal{W}$ is the set of all parameters for the network, and the set of training samples is $\{\underline{x}_p, \underline{d}_p\}_{p-1}^{P}$, then using y_j as an input to a cost function $\varepsilon(\cdot)$, whose minimization is achieved through a modified error backpropagation (Azimi 2018),

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

$$\varepsilon(\mathcal{W}) = \frac{1}{N}\sum_{p=1}^{P}\left\|\underline{d}_p - \underline{y}(\underline{x}_p;\mathcal{W})\right\|^2. \tag{11}$$

4.3 Residual Networks

For complicated data sets, network comparisons have shown that increased depth improves network accuracy (Simonyan and Zisserman 2015, Szegedy et al. 2015). Deep networks have more capacity for different level (low/mid/high) features (Zeiler and Fergus 2014), and the top performing networks on the ImageNet dataset have employed deep models (He et al. 2016)**.** The tradeoff is that the deeper the network becomes, the more the accuracy saturates and begins to quickly decline (He et al. 2016). To combat this, a group of researchers from Microsoft Research, He et al. (2016), have utilized the residual network architecture. A residual network architecture is very similar to a convolutional neural network, with one addition. Namely, every few convolutional layers, a short cut identity is included, Figure 6**,** which fits the layers to a residual mapping, instead of hoping that they would come to a desired mapping naturally (He et al. 2016).

4.4 Software Implementation

A major benefit to utilizing convolutional neural networks for hydrometeor classification based on high-resolution images is their (recent) widespread popularity for image processing in general. As a result of this popularity and widespread use, there are extensive software toolboxes available, both commercial and open source, with detailed walkthroughs for a variety of tasks and applications. Due to the automatic feature extraction inherent to their algorithm, CNNs can be applied and operated by non-experts, increasing their functionality as a preprocessing

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

frontend for big data tasks. The toolboxes and setup utilized in this study are described as follows.

The same network architecture is used for geometric classification and riming degree estimation and can be run in parallel. The network was implemented using MATLAB™ 2018, with the Deep Learning and Machine Learning Toolboxes at Colorado State University. The network architecture is the **ResNet-50**, which was chosen over **AlexNet** (Krizheysky et al. 2012) or **GoogleNet** (Szegedy et al. 2015) due to its balance between accuracy and speed (Mathworks 2018c) and has been pretrained on the ImageNet database (Russakovsky and Deng et al. 2015) to reduce the size of the training set data necessary to be an effective classifier. The more complicated the classification task, the larger the data set needed to extract relevant features (He et al. 2016). Geometric classification, however, is a common problem in image classification, and therefore it is an ideal candidate for pretraining. A network pretrained for image classification can reuse many of the features extracted and made task specific on a reduced training data set (del-Rio et al. 2018). Each class in the training set is limited to the smallest populated class, for instance, the Planar Crystal class is the least populated and has 290 images, this then sets an upper limit of 290 on all classes. This is done so overrepresentation does not occur in the training phase. The entire training data set is divided into two categories, “training” and “validation”. As is customary, 70% of the entire training data set is randomly selected and stored in the “training” category, with the remaining 30% saved for validation of the network performance. This separation reduces the likelihood of the network overfitting to data and allows for an accurate test for generalization. The training images are randomly reflected, translated and scaled within a defined range to improve the networks invariance to small changes (Murphy 2012). A technique known as Dropout (Srivastava et al. 2014) is employed on the pretrained

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

network weights to further reduce any overfitting that may occur. The network performance is determined by mean square and the loss function is calculated as stochastic gradient descent. Validation occurs every three iterations over 10 epochs (one epoch is a training phase where all training data is considered), although experimentation has shown that 6 epochs are enough to improve training time without loss of performance as is adopted for later tests. The learning rate for both networks is 0.0003.

5. Results and Discussion

This section presents and discusses the results of the geometric classification and riming degree estimation using the described method for classification of snowflakes based on images by a multi-angle snowflake camera by means of convolutional neural networks. The learning curves for each network are presented, along with an associated confusion matrix calculated from blind data to highlight the network's generalization (the ability for a network to classify new data).

5.1 Geometric Classification

The results of network training for geometric classification are shown in Figure 7. The training data set included ~1,450 images and training occurred over 900 iterations. The network achieves a mean accuracy of 93.4% with a loss function of ~0.2 and little variance between runs. Training time was 13 minutes and 23 seconds. This is very good accuracy given the size of the data set and the training time is reduced when limited to 6 epochs (from ~13 min to ~8 min). To test the generalization of the network, it is then used to classify ~400 snowflakes not included in training, the results of which are shown in Figure 8.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

The relevant information from the confusion matrix can be seen in the bottom row and right-most column. The bottom row is the percentage value of when the network was presented with an image and correctly classified it on a per class basis. The right-most column is out of class accuracy, i.e., how often the network confused an image within a class for something else. The confusion matrix shows that the network has some issues making decisions between planar crystal and aggregate snowflakes. This is consistent with established reasoning, as the planar crystal class is most visibly similar to aggregate snowflakes and was also the smallest represented class for training which limited the diversity of available samples. Increasing the number of samples should enhance the network's ability to differentiate between the two classes.

5.2 Riming Degree Estimation

With riming degree estimation, two networks were trained for comparison. One of the networks keeps the discrete five classes determined using Praz et al's (2017) [1,5] classification scheme (Table 1), while the other network removes the two classes labeled 2 and 4, relying instead on the posterior distribution, or "likelihood" estimation, to assign a continuous label to snowflakes that fall within the edge cases. Comparing the learning curves of the two networks (Figure 9), it is clear that classes 2 and 4 are difficult for the network to analyze. The result of training is a mean accuracy of 68.8% and a loss of ~1 for the network trained on five classes (Figure 7a) and a mean accuracy of 92.4% with a loss of <0.3 for the network with only three classes (Figure 7b). The second network is then tasked with classifying a data set that includes samples from all five classes. Results of the applied "likelihood" percentage, $R_{l,c\#}$, are shown in Figure 10. Here we show the network's capability to apply a riming degree estimation determined by features found in the edge case classes used in training. In Figure 10, examples of

Accepted for publication in Journal of Atmospheric and Oceanic Technology. DOI 10.1175/JTECH-D-19-0055.1.

individual degree estimations are depicted in the top-most row, while the network's capability to distinguish between snowflakes with similar characteristics is showcased on the bottom. An inherent benefit of this estimator is that it removes some of the guesswork involved when a human user classifies snowflakes individually. While the human processor may introduce error through fatigue or the result of an immediacy bias, a network will not suffer from these potential pitfalls and may catch misclassified images (Figure 11). Note that a larger, more diverse training data set will be required before more accurate and conclusive tests can be performed.

Developing a riming degree estimator relies almost entirely upon features that are unique to images of snowflakes. Best results will be derived from training sets where features unrelated to riming degree estimation are equally represented, so that they are effectively removed from the decision-making process. Without equal representation, the network can develop a bias based on which feature is more prevalent. An example of this bias shown in Figure 12 is a result of columnar crystal images being overrepresented in the "no riming" category. In Figure 12a, the estimator gives an accurate prediction indicated by the $R_{l,c\#}$ values, but the bias is evident in snowflakes in Figures 12b and 12c, resulting in a severe underestimation and overestimation, respectively. In the storms sampled, snowflakes demonstrating no degree of riming were relatively rare, which is why columnar crystal snowflakes compose the bulk of the data set. To achieve equal representation of unrelated snowflake features, a wider variety of storms from different seasons will need to be processed.

6. Conclusions

This paper has applied recent developments in machine learning to the problem of automatic winter hydrometeor classification. Utilizing convolutional neural networks, the task of

Accepted for publication in *Journal of Atmospheric and Oceanic Technology*. DOI 10.1175/JTECH-D-19-0055.1.

classifying snowflakes based on geometric characteristics and riming degree has been undertaken. Convolutional neural networks are ideal for image classification due to their efficient data handling, automatic feature extraction, versatility, and relative ease of application. A training set has been developed primarily from two winter precipitation events and consists of 1,450 snowflakes. Six geometric classes have been defined based on observable physical characteristics of the snowflakes. These classes are aggregate, columnar crystal, planar crystal, small particle, graupel, and a combination of columnar and planar (although the latter is discarded due to rarity until further data is processed). Geometric training has focused on individual snowflakes that are easily identifiable as members of a single class. The result of training is a network with 93.4% classification accuracy. This performance is sufficient to begin processing the many hours of recorded data from the MASCRAD project and begin growing the training data set for continued network development. Currently, our CNN is best equipped for small-batch processing (few hundred snowflakes per batch) due to the size of its training data set. By processing more data with diverse environmental conditions, the network will eventually be able to process bulk data that number in the thousands and tens of thousands, and more. With additional snowflake variety, geometric sub-classes and rarer classes may be introduced to the network, expanding from the five classes currently utilized.

The results obtained by the riming degree estimator using a CNN have shown promise. The training set is composed of images from all geometric classes separated (where applicable) into three categories: no riming, rimed, and graupel. The network has achieved 92% accuracy when estimating snowflakes that fall into these categories of riming. The probabilistic estimation that results at the output layer of the CNN has then been used to gauge where a snowflake falls within these three degrees of riming (classes 2 and 4 from Praz et al.). The high degree of

Accepted for publication in Journal of Atmospheric and Oceanic Technology. DOI 10.1175/JTECH-D-19-0055.1.

accuracy has been maintained by the network in determining whether a snowflake is more rimed or less rimed (belonging in class 2 or 4) but more sample data must be developed to increase precision of estimation and remove the influence of feature bias. Feature extraction requires a large number of images and features unique to snowflake images cannot be compensated for by pretraining on unrelated images. More diverse snowflakes will help remove any bias the network develops. For example, a method for mapping the estimation to fit Mosimann et al. (1994) can be developed as more output data becomes available. Finally, processing data from different seasons may lead to additional classification tasks, such as wet vs. dry snow, and may be considered in future applications of the CNN classification approach.

The classification network developed and presented in this paper will be used in the processing of MASCRAD data, but the architecture is suitable for any solid hydrometeor classification task and is suitable as a preprocessing frontend to any image-based particle recording instrument/device or system. The network is fitting for users with limited experience in image processing, machine learning or atmospheric research. Organized data by geometric and microphysical characteristics and accurate riming degree estimations will help further research into hydrometeor scattering. An example of future work is related to linking the microphysical characteristics of snowflakes to the scattering properties through 3D shape reconstruction and modeling (Kleinkort et al. 2017), followed by the realistic scattering computation (Chobanyan et al. 2015). Overall, this and many other applications of automatic CNN-based winter hydrometeor classification will potentially improve propagation models as higher frequencies continue to be explored, whether for use in remote sensing of hydrometeors, communication or other related fields.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

Acknowledgment

This work was supported in part by the National Science Foundation under Grant AGS-1344862 and in part by the National Aeronautics and Space Administration under PMM Science Grant NNX16AE43G.

References

Azimi, M. Lecture. ECE 656- Machine Learning and Adaptive Systems. Colorado State University, Fort Collins. Feb. 2nd – April 30th 2018.

Bang, W., K. Kim, L. GyuWon, M. Thurai, P. Kennedy, V.N. Bringi, and B. Notaros, 2016: Microphysical Characteristics Analysis of Three Heavy Snowfall Events from the MASCRAD Campain in Greeley, Colorado, USA, *URSI Asia-Pacific Radio Science Conference,* August 21-25, 2016, Seoul, Korea.

Besic, N., J. Figueras i Ventura, J. Grazioli, M. Gabella, U. Germann, and A. Berne 2016: Hydrometeor classification through statistical clustering of polarimetric radar measurements: a semi-supervised approach, *Atmos. Meas. Tech.*, **9**, 4425-4445, doi: 10.5194/amt-9-4425-2016.

Bringi, V. N., P. C. Kennedy, G.-J. Huang, C. Kleinkort, M. Thurai, and B. M. Notaros, 2017: Dual-polarized radar and surface observations of a winter graupel shower with negative Z_{dr} column," *J. Appl. Meteor. Climatol*., **56**, 455–470.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

Chandrasekar, V., R. Keranen, S. Lim, and D. Moisseev, 2013: Recent advances in classification of observations from dual polarization weather radars, *Atmos. Res.*, **119**, 97-111, doi: 10.1016/j.atmosres.2011.08.014.

Chobanyan E., N. J. Sekeljiic, A. B. Manic, M. M. Ilic, V. N. Bringi, and B. M. Notaroš, 2015: Efficient and Accurate Computational Electromagnetics Approach to Precipitation Particle Scattering Analysis Based on Higher Order Method of Moments Integral-Equation Modeling. *J. Atmos. Oceanic Technol.*, **32**, 1745–1758.

Collobert, R., and J. Weston, 2008: A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning, *Proceedings of the 25th international conference on Machine learning, 160-167* doi:10.1145/1390156.1390177

Feind, R. E., 2006: Comparison of three classification methodologies for 2D probe hydrometeor images obtained from the armored T-28 aircraft, PhD thesis, South Dakota School of Mines and Technology.

Garrett, T. J., C. Fallgatter, K. Shkurko, and D. Howlett, 2012: Fall speed measurement and high-speed resolution muli-angle photography of hydrometeors in free fall, *Atmos. Meas. Tech.*, **5**, 2625-2633, doi:10.5194/amt-5-2625-2012.

Goodfellow, I., Y. Bengio, and A. Courville, 2016: Deep Learning, *The MIT Press* 322-359

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

Grazioli, J., D. Tuia, S. Monhart, M. Schneebeli, T. Raupach, and A. Berne, 2014: Hydrometeor classification from two-dimensional video disdrometeor data, *Atmos. Meas. Tech.*, **7**, 2869-2882, doi:10.5194/amt-7-2869-2014.

Haykin, S., 2009: Neural Networks and Learning Machines Third Edition, *Pearson Education Inc.* 123-217

He, K., X. Zhang, S. Ren, and J. Sun, 2016: Deep Residual Learning for Image Recognition, *CVPR*.

Kennedy, P., M. Thurai, C. Praz, V. N. Bringi, A. Berne, and B. M. Notaros, 2018: Variations in Snow Crystal Riming and Z_{DR}: A Case Analysis. *J. Appl. Meteor. Climatol.*, **57**, 695–707.

Kim, M. J., 2006: Single scattering parameters of randomly oriented snow particles at microwave frequencies, *Journal of Geophysical Research*, col. 111, D14201, doi:10.1029/2005JD006892.

Kleinkort, C., G.-J. Huang, V. N. Bringi, and B. M. Notaros, 2017: Visual Hull Method for Realistic 3D Particle Shape Reconstruction Based on High-Resolution Photographs of Snowflakes in Free Fall from Multiple Views. *J. Atmos. Oceanic Technol.*, **34**, 679–702.

Korolev, A. and B. Sussman, 2000. A technique for habit classification of cloud particles, *J. Atmos. Ocean. Tech.*, 17, 1048-1057.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology*. DOI 10.1175/JTECH-D-19-0055.1.

Krizhevsky, A., I. Sutskever, and G.E. Hinton, 2012: Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pp 1097-1105.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner, 1998: Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, **86**(11), 2278-2324.

Libbrecht, K. G., 2017: Physical Dynamics of Ice Crystal Growth, *Annu. Rev. Mater. 2017.* **47**, 271-295, doi: 10.1146/annurev-matsci-070616.

Magono, C. and C.W. Lee, 1966: Meteorological classification of natural snow crystals, *J. Fac. Sci.*, Hokkaido Univ., Series VII, 2, 321-335.

Mathworks, 2018a: Classify Image Using GoogLeNet. August 2018. https://www.mathworks.com/help/deeplearning/examples/classify-image-using-googlenet.html

Mathworks, 2018b: Train Deep Learning Network to Classify New Images. August 2018. https://www.mathworks.com/help/deeplearning/examples/train-deep-learning-network-to-classify-new-images.html

Accepted for publication in *Journal of Atmospheric and Oceanic Technology*. DOI 10.1175/JTECH-D-19-0055.1.

Mathworks, 2018c: Pretrained Convolutional Neural Networks. August 2018. https://www.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html#mw_45a8c0b2-26fa-48e9-905a-a7ed7b87bfc8

MATLAB and Machine Learning and Deep Learning Toolbox Release 2018, The MathWorks, Inc., Natick, Massachusetts, United States https://www.mathworks.com/solutions/deep-learning.html?s_tid=hp_brand_deeplearning

Minar, M.N., and J. Naher, 2018: Recent Advances in Deep Learning: An Overview. 10.13140/RG.2.2.24831.10403.

Mosimann, L., E. Weingartner, and A. Waldvogel, 1994: An analysis of accreted drop sizes and mass on rimed snow crystals, *J. Atmos. Sci*, **51**, 1548-1558.

Murphy, K. P., 2012: Machine Learning, A Probabilistic Perspective, *The MIT Press*, 566-572

Newman, A. J., P. A. Kucera, and L. F. Bliven, 2009: Presenting the Snowflake Video Imager (SVI). *J. Atmos. Oceanic Technol.*, **26**, 167–179.

Notaroš, B. M., V. N. Bringi, C. Kleinkort, P. Kennedy, G.-J. Huang, M. Thurai, A. J. Newman, W. Bang, and G. Lee, 2016: Accurate Characterization of Winter Precipitation Using Multi-Angle Snowflake Camera, Visual Hull, Advanced Scattering Methods and Polarimetric Radar.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

invited paper, Special Issue on Advances in Clouds and Precipitation, *Atmosphere*, **7**, no. 6, 81–111.

O'Shea, T. J., J. Corgan, and T.C. Clancy, 2016: Convolutional Radio Modulation Recognition Networks. In: Jayne C., Iliadis L. (eds) *Engineering Applications of Neural Networks. EANN* 2016. *Communications in Computer and Information Science*, **629**. Springer, Cham. doi: 10.1007/978-3-319-44188-7_16.

Praz, C., R. Yves-Alain, and A. Berne, 2017: Solid hydrometeor classification and riming degree estimation from pictures collected with a Multi-Angle Snowflake Camera, *Atmos. Meas. Tech.*, **10**, 1335-1357.

del-Rio, F., P. Messina, V. Dominguez, and D. Parra, 2018: Do Better ImageNet Models Transfer Better... for Image Recommendation? *CoRR, abs/1807.09870*.

Romanuke, V., 2017: Appropriate Number of Standard 2x2 Max Pooling Layers and Their Allocation in Convolutional Neural Networks for Diverse and Heterogeneous Datasets, *Info. Tech. and Mgmt. Sci.*, Sciendo, **20**(1), pages 12-19, December.

Russakovsky, O.,*, J. Deng,*, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, and L. Fei-Fei, (* = equal contribution) ImageNet Large Scale Visual Recognition Challenge. *IJCV*, 2015.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology*. DOI 10.1175/JTECH-D-19-0055.1.

Schönhuber, M., Lammer, G., and Randeu, W., 2008: The 2D video disdrometer. In Precipitation: Advances in Measurement, Estimation and Prediction; Michaelides, S., Ed.; *Springer*: Berlin, Germany, 3–31.

Simonyan, K., and A. Zisserman, (2015):Very deep convolutional networks for large-scale image recognition. *ICLR*, 2015.

Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, 2014: Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *JMLR. 15 1929-1958*

Straka, J., D. S. Zrnić, and A. V. Ryzhkov, 2000: Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of Relations. *J. Appl. Meteor.*, **39**, 1341–1372.

Svozil, D., V. Kvasnička, and J. Pospíchal, 1997: Introduction to multi-layer feed-forward networks, *Chemometrics and Intelligent Laboratory Systems* **39** (1997) 43-62

Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, 2015: Going deeper with convolutions. *CVPR.*

Torrey, L., and J. Shavlik, 2009: Transfer Learning, *Handbook of Research on Machine Learning Applications*, by IGI Global.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology*. DOI 10.1175/JTECH-D-19-0055.1.

Tyynelä, J., J. Leinonen, D. Moisseev, and T. Nousiainen, 2011: Radar Backscattering from Snowflakes: Comparison of Fractal, Aggregate, and Soft Spheroid Models. *J. Atmos. Oceanic Technol.*, **28**, 1365–1372.

Zeiler, M.D. and R. Fergus, 2014: Visualizing and understanding convolutional neural networks. *ECCV*.

Zhang, G., S. Luchs, A. Ryzhkov, M. Xue, L. Ryzhkova, and Q. Cao, 2011: Winter Precipitation Microphysics Characterized by Polarimetric Radar and Video Disdrometer Observations in Central Oklahoma. *J. Appl. Meteor. Climatol.*, **50**, 1558–1570.

Tables

Table 1. Physical descriptions characterizing the degree of riming, R_d, on a given snowflake, with numerical representations as described by Mosimann et al. (1994), in column 2. Column 1 contains degree estimations, R_c, utilized by Praz et al. (2017). Column 3 gives the probability estimates, R_l, used in this paper.

$R_c \in [0,1]$	$R_d \in [1,5]$	$R_l \in [1,5]$	Coverage of the surface	Description
0	1 (none)	1.0 ($R_{l,c1} \geq 99\%$)	0%	No cloud droplets on the surface. Snowflakes are detailed and delicate in appearance.
0.15	2 (rimed)	to 2.99 ($R_{l,c1}$, $R_{l,c3}$ $> R_{l,5}$)	$\simeq$ 50%	Up to half of the surface is covered with cloud droplets. There may be delicate features, but some clumping has occurred.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

0.5	3 (densely rimed)	3.0 ($R_{l,c3} \geq 99\%$)	$\simeq 100\%$	The entire snowflake is covered with cloud droplets, but the general shape is conserved.
0.85	4 (graupel-like)	3.01 to 4.99 ($R_{l,c5}$, $R_{l,c3} > R_{l,c1}$)	$> 100\%$	The entire snowflake is heavily covered with cloud droplets, to the point where the original shape is barely recognizable
1.0	5 (graupel)	5.0 ($R_{l,c5} \geq 99\%$)	$\gg 100\%$	The entire snowflake is heavily covered with cloud droplets. Original shape is no longer distinguishable and has entered the class of graupel.

Figure Caption List

Figure 1. MASCRAD Snow Field Site at Easton Valley Airport, near Greely, Colorado, under the umbrella of CSU-CHILL Radar. MASC (top right), along with other surface instrumentation, is contained in the 2/3-scaled DFIR.

Figure 2. Multi-Angle Snowflake Camera (MASC), with three cameras in the horizontal plane for capturing high-resolution photographs of winter precipitation (top left); The CSU MASC system has been modified to included two cameras situated on an elevated plane, poised at 55° above the horizon to aid in the visual hull 3D particle shape reconstruction (top right); Example images from each camera, the first three columns are from the horizontal cameras and the final two are from the lower-resolution raised cameras (bottom two rows, from left to right).

Figure 3. Examples of MASC images characterizing the geometric classes (top row) and riming degree estimations (bottom row).

Accepted for publication in *Journal of Atmospheric and Oceanic Technology*. DOI 10.1175/JTECH-D-19-0055.1.

Figure 4. (a) Feed Forward Network architecture demonstrating how each neuron is connected to every neuron in the previous layer. (b) Algorithm for feed forward learning process within the network.

Figure 5. Highlighting the key steps within a typical convolutional neural network: The input layer is scanned with a $[i,j]$ weight matrix to create the feature map; The feature map is averaged or maxed to increase network efficiency in the pooling layer; Decision making occurs at the final fully connected layer.

Figure 6. Additional Residual step introduced to a typical neural network architecture.

Figure 7. A mean average is validated every three iterations (black dots) over 10 epochs with a mean average of 93.4% (top), the same parameters are shared by the loss function evaluation with a final value of ~0.2 (bottom).

Figure 8. The network was tasked with classifying 395 snowflakes to test generalization. The left axis is what the network labeled the input data as, and the bottom axis is what the input data was. Green boxes represent correctly classified images, with total number in bold and percentage of total immediately below. Green text represents a correct classification, while the red percentage is the misclassification of the network. The overall network accuracy is shown in the bottom right corner.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

Figure 9. (a) A network was trained using classes composed of images representing the five discrete riming degree values [1,5] described in Table 1; (b) A separate network was trained using the same image data, only removing the images comprising classes 2 and 4.

Figure 10. The values represent the output of the riming degree estimation network: the top row highlights the network's range, while the bottom mostly demonstrates its ability to differentiate between similar snowflakes.

Figure 11. Utilizing a neural network for riming degree estimation has the advantage of quantified data driven decision making.

Figure 12. (a) An image of a correctly labeled image by the Riming degree estimator. (b) The network underestimates the degree of riming in the image due to feature bias present in the "no riming" class. Expected results for this image should be a higher value for $R_{l,c3}$. (c) The network overestimates the degree of riming due to feature bias from the "graupel" class.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

Figures

Figure 1. MASCRAD Snow Field Site at Easton Valley Airport, near Greely, Colorado, under the umbrella of CSU-CHILL Radar. MASC (top right), along with other surface instrumentation, is contained in the 2/3-scaled DFIR.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

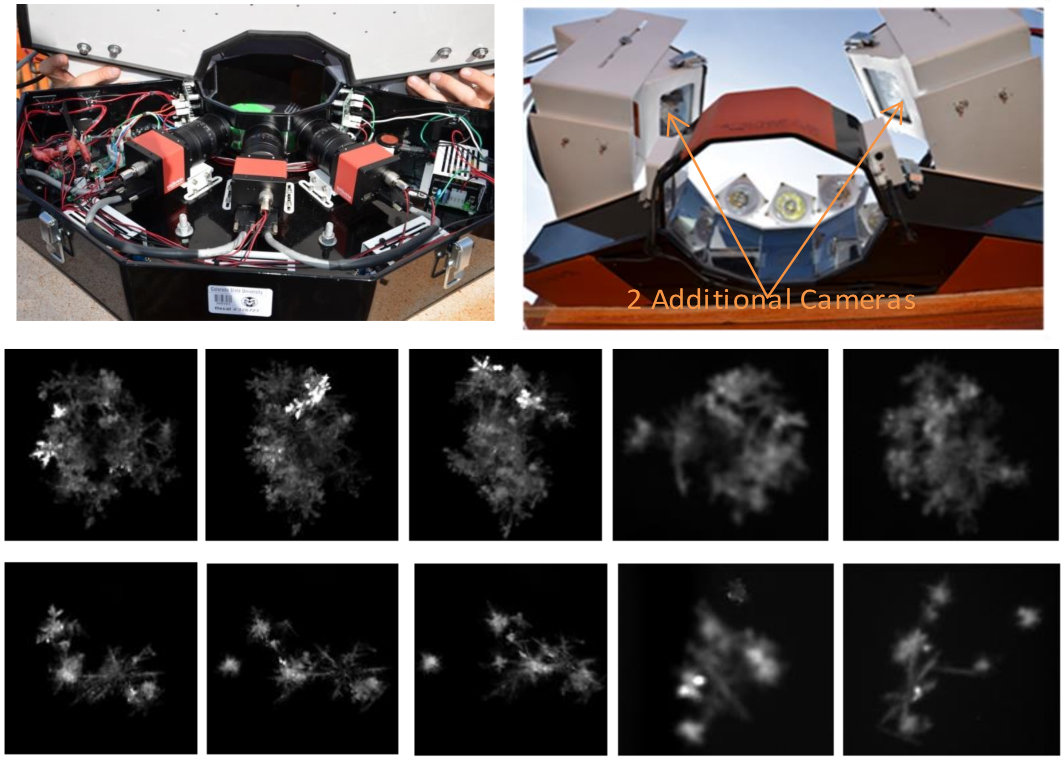

Figure 2. Multi-Angle Snowflake Camera (MASC), with three cameras in the horizontal plane for capturing high-resolution photographs of winter precipitation (top left); The CSU MASC system has been modified to included two cameras situated on an elevated plane, poised at 55° above the horizon to aid in the visual hull 3D particle shape reconstruction (top right); Example images from each camera, the first three columns are from the horizontal cameras and the final two are from the lower-resolution raised cameras (bottom two rows, from left to right).

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

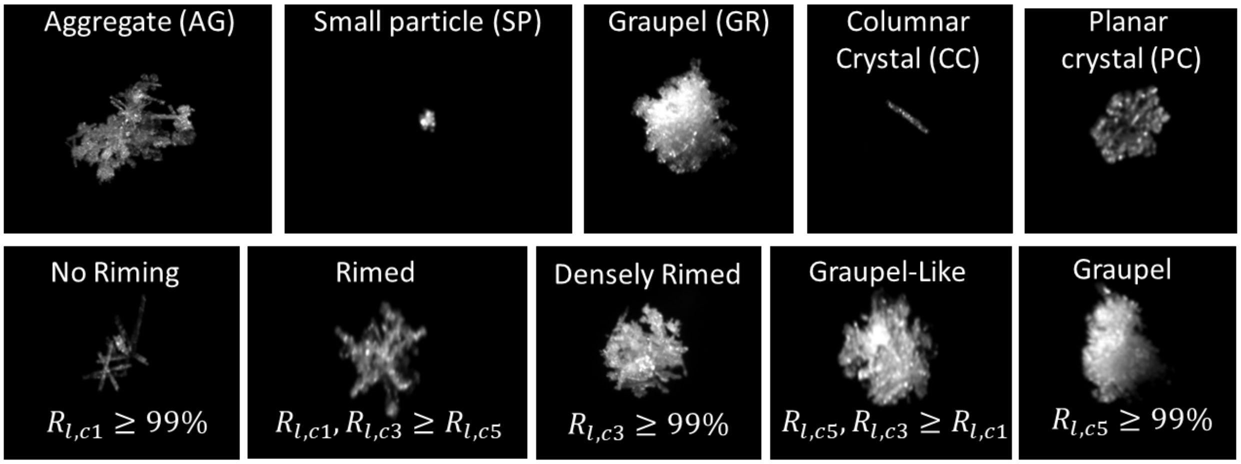

Figure 3. Examples of MASC images characterizing the geometric classes (top row) and riming degree estimations (bottom row).

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

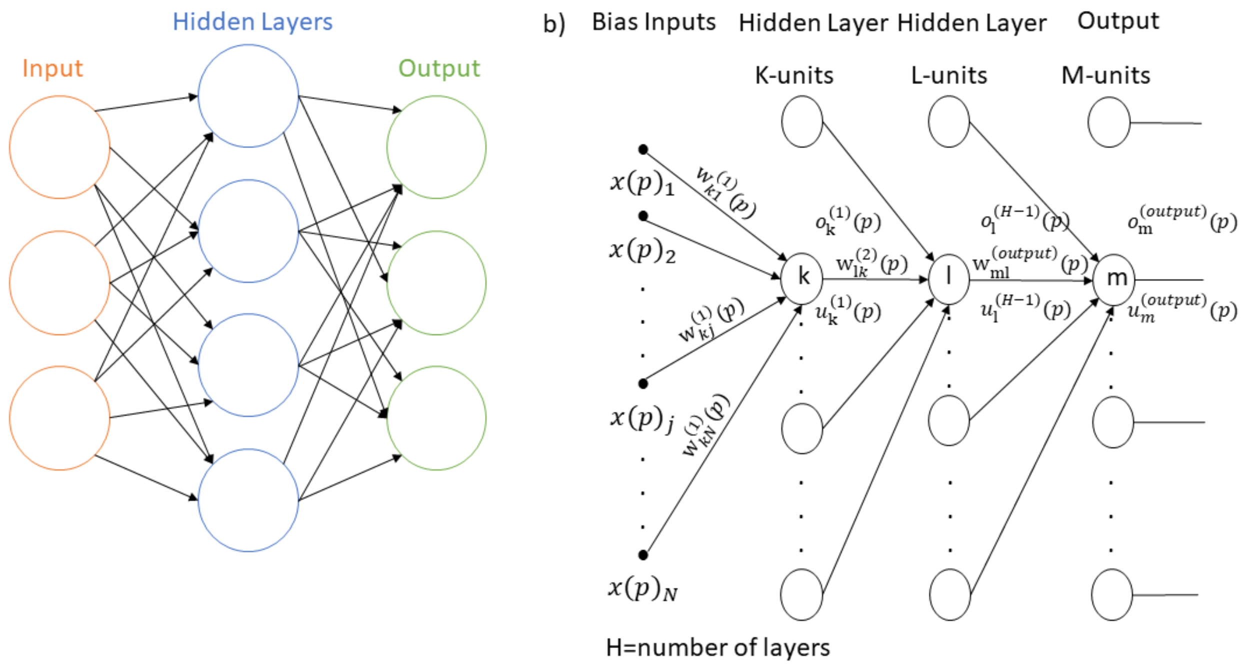

Figure 4. (a) Feed Forward Network architecture demonstrating how each neuron is connected to every neuron in the previous layer. (b) Algorithm for feed forward learning process within the network.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

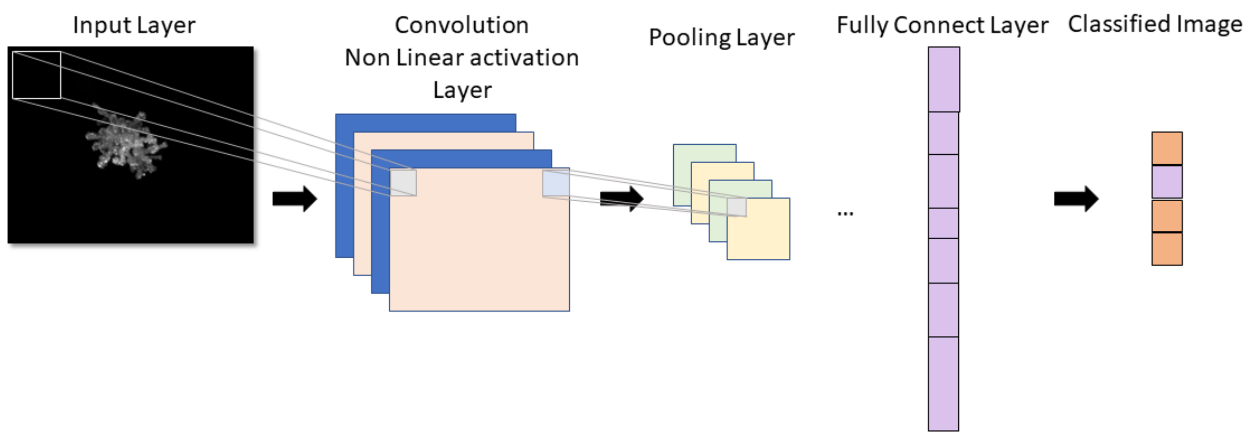

Figure 5. Highlighting the key steps within a typical convolutional neural network: The input layer is scanned with a [*i*,*j*] weight matrix to create the feature map; The feature map is averaged or maxed to increase network efficiency in the pooling layer; Decision making occurs at the final fully connected layer.

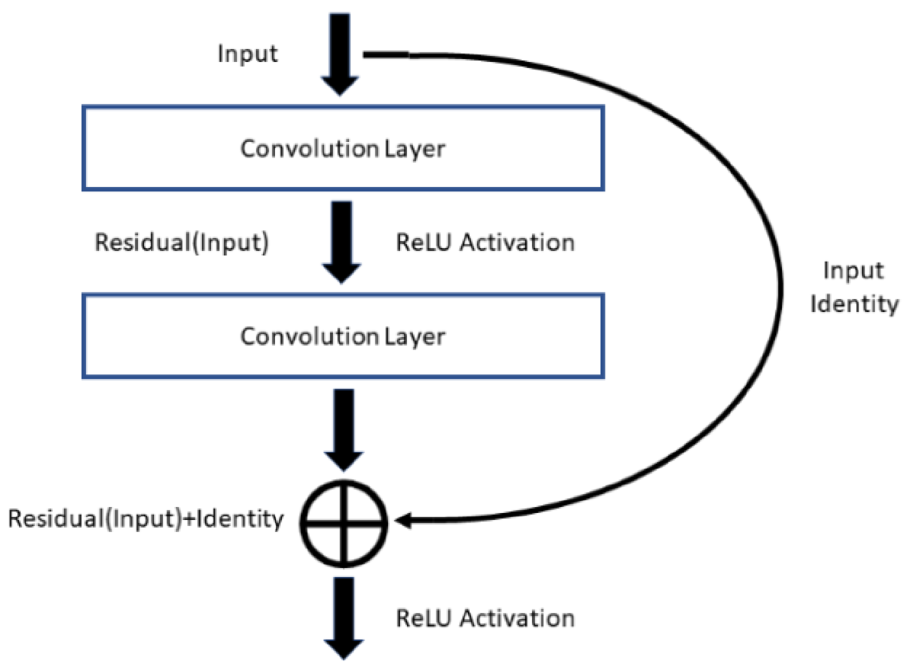

Figure 6. Additional Residual step introduced to a typical neural network architecture.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

Figure 7. A mean average is validated every three iterations (black dots) over 10 epochs with a mean average of 93.4% (top), the same parameters are shared by the loss function evaluation with a final value of ~0.2 (bottom).

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

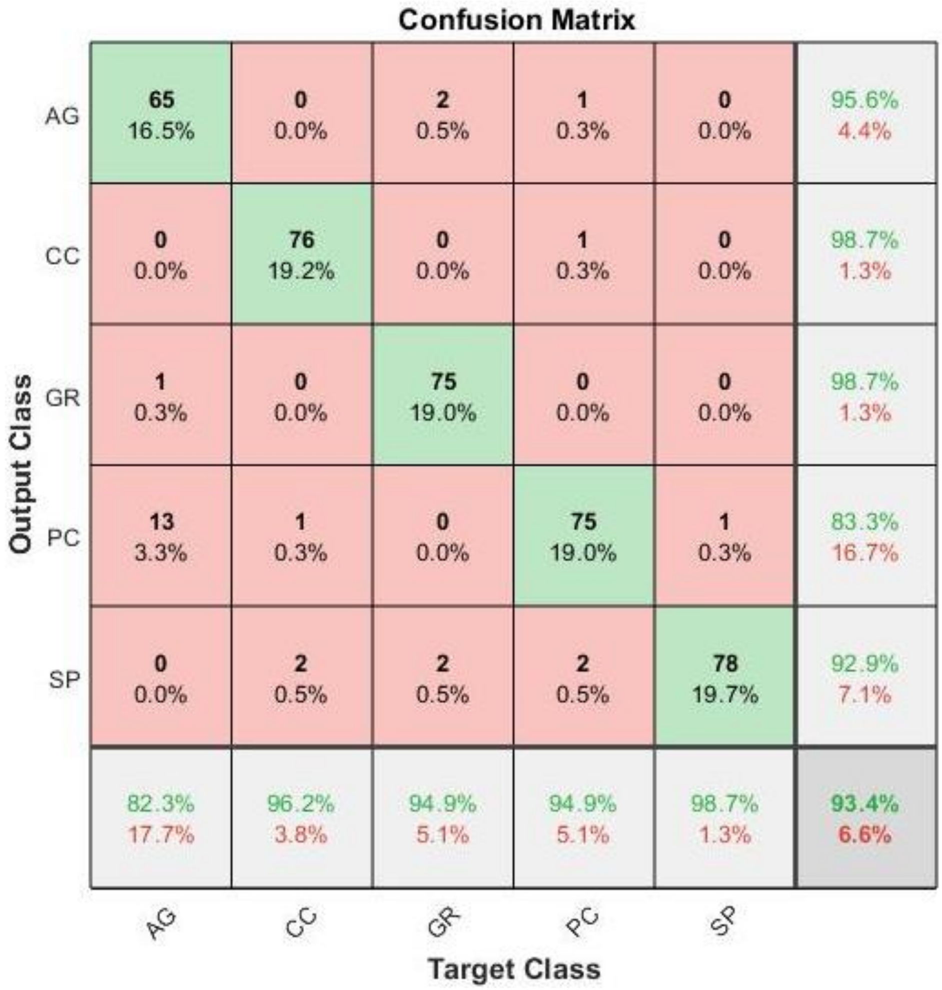

Figure 8. The network was tasked with classifying 395 snowflakes to test generalization. The left axis is what the network labeled the input data as, and the bottom axis is what the input data was. Green boxes represent correctly classified images, with total number in bold and percentage of total immediately below. Green text represents a correct classification, while the red percentage is the misclassification of the network. The overall network accuracy is shown in the bottom right corner.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

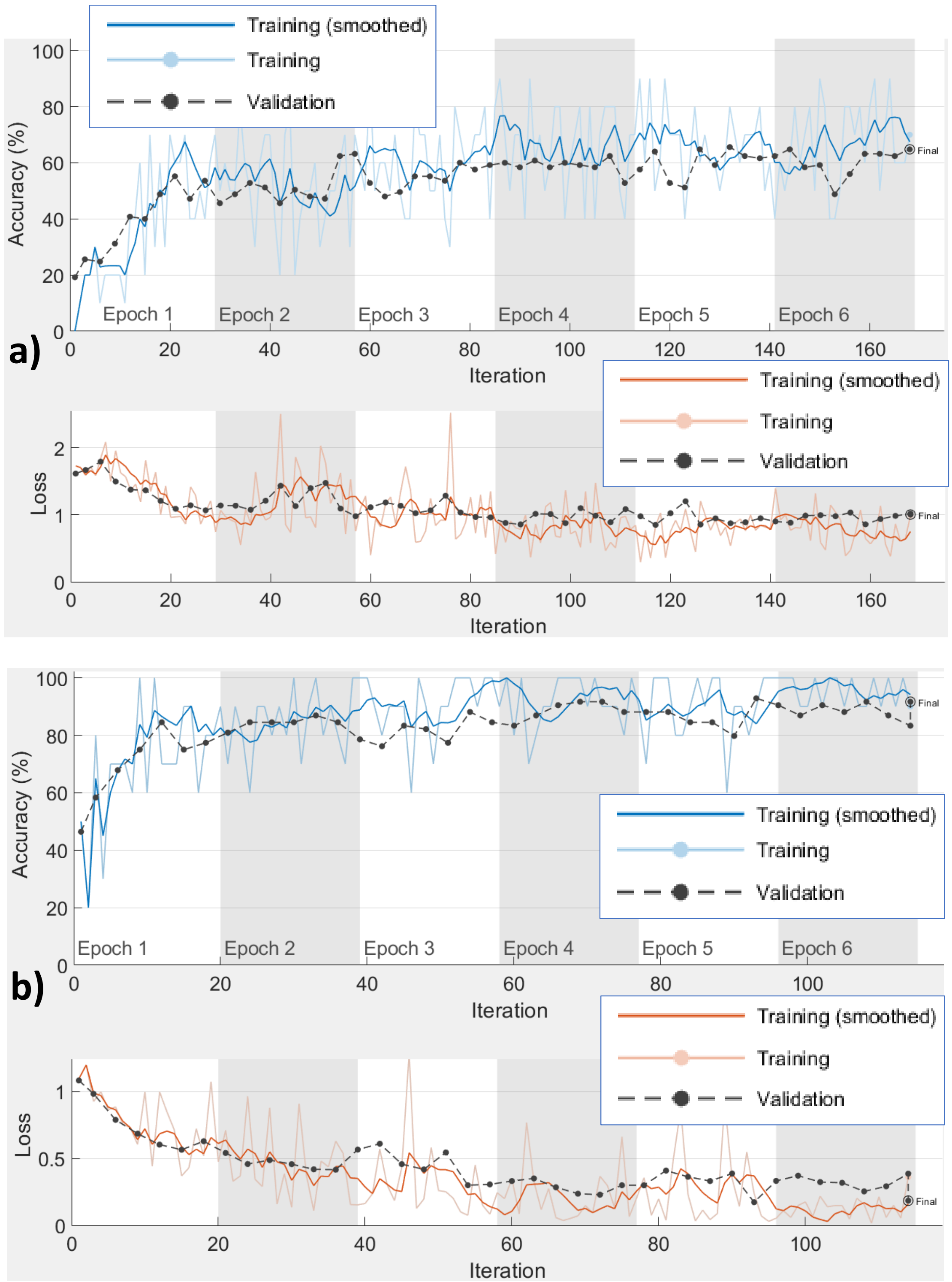

Figure 9. (a) A network was trained using classes composed of images representing the five discrete riming degree values [1,5] described in Table 1; (b) A separate network was trained using the same image data, only removing the images comprising classes 2 and 4.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

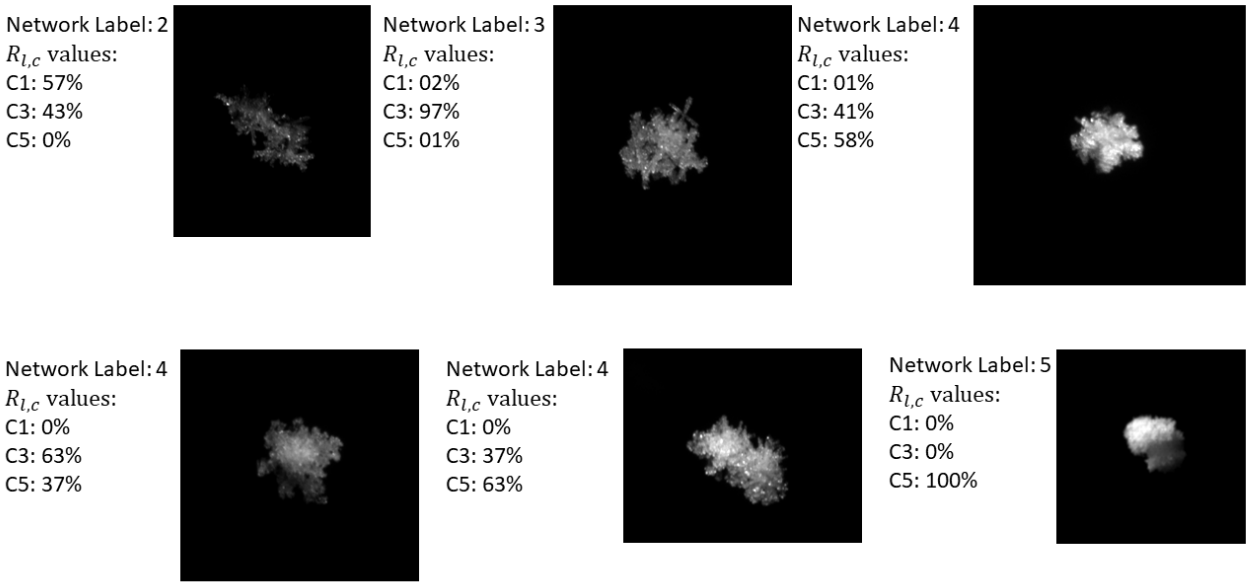

Figure 10. The values represent the output of the riming degree estimation network: the top row highlights the network's range, while the bottom mostly demonstrates its ability to differentiate between similar snowflakes.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.

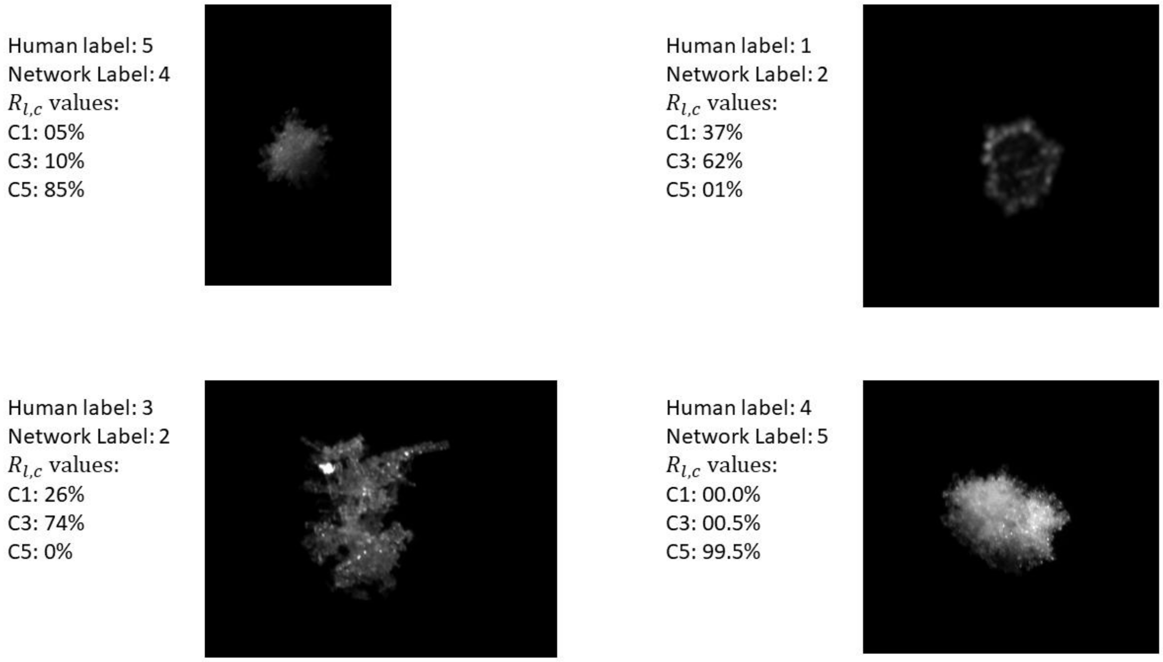

Figure 11. Utilizing a neural network for riming degree estimation has the advantage of quantified data driven decision making.

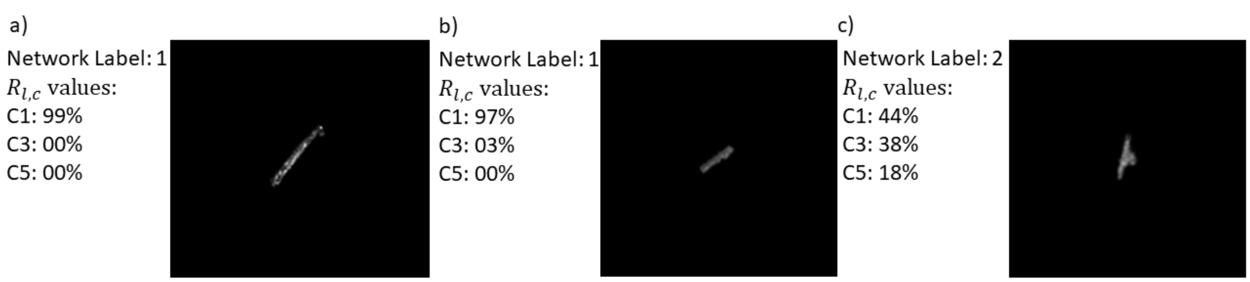

Figure 12. (a) An image of a correctly labeled image by the Riming degree estimator. (b) The network underestimates the degree of riming in the image due to feature bias present in the "no riming" class. Expected results for this image should be a higher value for $R_{l,c3}$. (c) The network overestimates the degree of riming due to feature bias from the "graupel" class.

Accepted for publication in *Journal of Atmospheric and Oceanic Technology.* DOI 10.1175/JTECH-D-19-0055.1.